REPUBLIKA E SHQIPËRISË
UNIVERSITETI POLITEKNIK I TIRANËS
FAKULTETI I TEKNOLOGJISË SË INFORMACIONIT
DEPARTAMENTI I INXHINIERISË INFORMATIKE

ALBA HAVERIKU

PËR MARRJEN E GRADËS
"DOKTOR"
NË: "TEKNOLOGJITË E INFORMACIONIT DHE KOMUNIKIMIT"
DREJTIMI "INXHINIERI INFORMATIKE"

DISERTACION

MODELE TË INTELIGJENCËS ARTIFICIALE NË PËRPUNIMIN E TEKSTEVE NË GJUHËN SHQIPE PËRMES INTEGRIMIT TË TË DHËNAVE NGA GJURMIMI I SYRIT DHE I MAUSIT

Udhëheqës shkencor
Prof. Dr. Elinda MEÇE

TIRANË, 2026

MODELE TË INTELIGJENCËS ARTIFICIALE NË PËRPUNIMIN E TEKSTEVE NË GJUHËN SHQIPE PËRMES INTEGRIMIT TË TË DHËNAVE NGA GJURMIMI I SYRIT DHE I MAUSIT

Disertacioni

i paraqitur në Universitetin Politeknik të Tiranës

për marrjen e gradës

"Doktor"

në:

"Teknologjitë e Informacionit dhe Komunikimit"

Drejtimi "Inxhinieri Informatike"

Nga

Znj. Alba Haveriku

TIRANË, 2026

JURIA PËR VLERËSIMIN E DISERTACIONIT PËR FITIMIN E GRADËS SHKENCORE "DOKTOR"

Miratuar

Me vendimin e Këshillit të Profesorëve të FTI-së Nr: 7, datë: 02.04.2026.

Kryetari i Jurisë (Oponent):	Prof. Asoc. Hakik PACI
Anëtar i Jurisë:	Prof. Asoc. Enida SHEME
Anëtar i Jurisë:	Prof. Asoc. Dorian MINAROLLI
Anëtar i Jurisë:	Prof. Dr. Alda KIKA
Anëtar i Jurisë (Oponent):	Prof. Dr. Ana KTONA

Dekan i Fakultetit të Teknologjisë së Informacionit

Prof. Dr. Elinda MEÇE

TABELA E PËRMBAJTJES

ABSTRAKT

Përpunimi i gjuhës natyrore (PGJN) (angl. *Natural Language Processing* – NLP) përbën një nga fushat kryesore të kërkimit dhe zhvillimit në Inteligjencën Artificiale (IA). Ekzistenca e mijëra gjuhëve në botë e vështirëson ndërtimin e modeleve kompjuterike gjithëpërfshirëse, pasi secila gjuhë është e lidhur me rregulla specifike gjuhësore, me raste përjashtimore dhe variacione kontekstuale të cilat kërkojnë ndërtimin dhe përmirësimin e vazhdueshëm të sistemeve inteligjente. Zhvillimet e fundit të lidhura me përdorimin në shkallë të gjerë të modeleve të mëdha gjuhësore (angl. *Large Language Models – LLMs*) kanë sjellë ndryshime dhe zhvillime në fushën e përpunimit të gjuhës natyrore. Megjithatë, këto modele kërkojnë bashkësi të mëdha të dhënash, të cilat nuk janë gjithmonë të disponueshme, sidomos në rastin e gjuhëve me burime të pakta (angl. *Low-Resource Languages – LRLs*).

Gjatë dekadës së fundit, vihet re rritje e numrit të punimeve shkencore të cilat studiojnë përdorimin e të dhënave konjitive të mbledhura nga eksperimentet psikolinguistike për të përmirësuar performancën e modeleve ekzistuese gjuhësore. Kërkimet në psikolinguistikë dhe neurolinguistikë ofrojnë një bazë të rëndësishme teorike dhe metodologjike për zhvillimin e modeleve kompjuterike të avancuara të përpunimit të gjuhës. Edhe në rastin e PGJN-së, të dhënat konjitive janë provuar të ofrojnë ndihmë në ndërtimin dhe përmirësimin e modeleve të cilat kanë aftësinë të përshtaten me struktura dhe kontekste të ndryshme gjuhësore. Të dhënat nga eksperimentet me gjurmimin e syrit gjatë leximit dhe alternativa të tjera me kosto të ulët, si gjurmimi i lëvizjeve të mausit, janë përdorur me sukses për etiketime automatike të pjesëve të ligjëratës, identifikimin e kategorive sintaksore, njohjen e entiteteve emërore, zbulimin e gabimeve gramatikore etj.

Kërkimet e PGJN-së për gjuhën shqipe, si një gjuhë me burime të pakta, janë rritur gjatë viteve të fundit, por hasin vështirësi të konsiderueshme që lidhen me mungesën e korpuseve të mëdha të etiketuar dhe me mungesën e standardeve për përpunimin e tekstit. Ndërkohë, fusha e gjurmimit të syrit mbetet ende pak e zhvilluar për studimin e sjelljeve gjatë leximit për gjuhën shqipe. Ekziston një mungesë e theksuar e studimeve psikolinguistike të lidhura me gjuhën shqipe, gjë që kufizon përfshirjen e saj në studime krahasuese ndërgjuhësore.

Kontributet kryesore të këtij punimi qëndrojnë në vlerësimin e situatës aktuale për zhvillimin e eksperimenteve psikolinguistike gjatë leximit dhe vlerësimin e metodave për integrimin e të dhënave konjitive në modelet e inteligjencës artificiale. Në këtë punim kemi paraqitur një rishikim sistematik të literaturës së dekadës së fundit, ku analizohen studimet më të rëndësishme që trajtojnë ndërthurjen e dy fushave: PGJN-së dhe psikolinguistikës. Kjo analizë ofron një pasqyrë të zhvillimeve të fundit në këtë drejtim duke theksuar

tendencat kryesore kërkimore, metodat e përdorura dhe sfidat e ndërlidhura në integrimin e të dhënave konjitive në modelet e inteligjencës artificiale.

Kontribut i rëndësishëm i këtij punimi është krijimi i dy korpuseve të para të llojit të tyre për gjuhën shqipe, një korpus i krijuar nga zhvillimi i eksperimenteve të gjurmimit të mausit dhe një korpus nga gjurmimi i syrit gjatë leximit. Këto dy korpuse përbëjnë një burim të vlefshëm të të dhënave për studimin e proceseve konjitive gjatë leximit në gjuhën shqipe. Përtej krijimit të korpuseve në këtë punim janë vlerësuar metoda për integrimin e të dhënave nga eksperimentet e gjurmimit të syrit dhe të mausit në detyra të përpunimit të gjuhë natyrore si: etiketimi i pjesëve të ligjëratës, analiza e lidhjeve sintaksore dhe identifikimi i fjalëve kyçe. Rezultatet eksperimentale të arritura nga integrimi i të dhënave të gjurmimit të syrit dhe të mausit në modele të inteligjencës artificiale tregojnë se ky tip të dhënash mund të kontribuojë në përmirësimin e performancës së modeleve në disa detyra të ndryshme.

Ky studim kontribuon në zgjerimin e kërkimeve ndërdisiplinore mes PGJN-së, inteligjencës artificiale dhe psikolinguistikës, duke ofruar një bazë të re dhe të aksesueshme të të dhënave, një metodologji eksperimentale dhe analizë të rezultateve. Eksperimentimi me dy metodologji eksperimentale ofroi mundësinë e krahasimit të alternativave me kosto të lartë (gjurmimi i syrit përmes pajisjes EyeLink Portable Duo) dhe me kosto të ulët (gjurmimi i kursorit të mausit). Analiza krahasuese e metrikave të leximit vërtetoi se ekziston një korrelacion pozitiv, sidomos për metrikën e kohëzgjatjes totale të leximit për fjalë.

Gjithashtu, përfshirja e gjuhës shqipe në korpuse paralele shumë gjuhësore është një kontribut shumë i vlefshëm pasi lejon studimin krahasues të strukturave gjuhësore dhe analizën e tyre në raport me gjuhë të tjera. Në këtë mënyrë, shqipja mund të analizohet jo vetëm në vetvete, por edhe në një perspektivë tipologjike dhe ndërgjuhësore. Vlera e metrikave të vlerësimit të modeleve të inteligjencës artificiale në eksperimentet e zhvilluara gjatë këtij punimi rriten me ritëm të ngjashëm me punimet në gjuhë të tjera të identifikuara gjatë rishikimit të literaturës, duke e vendosur shqipen në nivel krahasues me to. Rezultatet e arritura theksojnë ndikimin e arkitekturës së përdorur dhe konfigurimeve të ndryshme eksperimentale në parashikimin e etiketave të ndryshme gjuhësore. Së fundmi, integrimi i këtyre të dhënave për përmirësimin e modeleve të PGJN-së hapin mundësi për zhvillime të mëtejshme në përpunimin kompjuterik të gjuhës shqipe dhe në ndërtimin e burimeve të dobishme për komunitetin shkencor.

Fjalë kyçe: gjurmimi i syrit, gjurmimi i mausit, përpunimi i gjuhëve natyrore, etiketimi i pjesëve të ligjëratës, etiketimi i lidhjeve sintaksore, identifikimi i fjalëve kyçe, gjuha shqipe.

ABSTRACT

Natural Language Processing (NLP) tasks constitute one of the main fields of study and research in Artificial Intelligence (AI) applications. The existence of thousands of different languages in the world is one of the key issues regarding the creation of general computational models, since each language is composed of different linguistic rules, exceptional cases and contextual variations, all of which require continuous development and refinement of intelligent systems. The adoption of Large Language Models (LLMs) has significantly impacted the field of NLP. However, these models require large-scale datasets which are not always available, particularly in the case of low-resource languages.

In the past decade, there has been an increase in scientific studies investigating the use of cognitive data collected during psycholinguistic experiments to improve the performance of existing language models. The research done in the fields of neurolingustic and psycholinguistic can offer a solid theoritical and methodological base for the development of advanced computational models. In the context of NLP, cognitive data has proven to be valuable in the development and improvement of models that adapt to new linguistic structures and contexts. The data obtained from eye-tracking experiments during reading, as well as other low-cost tracking methods, such as mouse tracking, have been successfully used to detect part of speech, syntactic categories, named entity recognition, the detection of grammatical errors and other related tasks.

NLP related research for the Albanian language as a low-resource language has increased over the last years. However, it continues to face a lot of challenges and limitations, particularly the lack of large annotated corpuses and the lack of standards on text processing. Meanwhile, regarding the field of eye-tracking research, the study of the Albanian language is underexplored. There is a lack of psycholinguistic studies linked with Albanian, which limits its inclusion in crosslingual studies.

The main contributions of this work lie in the assessment of the current state of psycholinguistic reading experiments and in the evaluation of methods integrating cognitive data into artificial intelligence models. This study presents a systematic review identifying and analyzing the most important research works of the past decade that study the intersection of NLP and psycholinguistics. This analysis provides an overview of recent developments in this research field, focusing on the main research methods, tools and challenges related to the integration of cognitive data in artificial intelligence models.

The most important contribution of this research is the creation of the first two corpora of their kind for the Albanian language. The first one, a corpus collected from mouse tracking experiments and the second one composed of the data collected from eye-tracking experiments. These two corpora provide a good resource for studying cognitive processes

while reading in the Albanian language. Furthermore, in this study we evaluate the inclusion of the data collected from eye-tracking and mouse tracking experiments for different NLP tasks, such as the identification of part of speech tags, dependency parsing and keywords extraction. The integration of the data collected from eye-tracking and mouse tracking experiments in artificial intelligence models proves that this type of data can contribute in the improvement of the performance of the models in various NLP subtasks.

This study contributes to the expansion of interdisciplinary research in NLP, artificial intelligence and psycholinguistics, offering a database of openly accessible data, a clear experimental methodology and results' analysis. Analysing two experimental methodologies, provided the possibility to compare higher accuracy alternatives (eye-tracking with the EyeLink Portable Duo device) and low cost alternatives (mouse tracking). The comparative analysis of reading metrics confirmed the existence of a positive correlation, especially for the total reading time per word metric.

Overmore, the inclusion of the Albanian language in a multilingual parallel corpora presents a valuable contribution that facilitates the study and comparison of Albanian linguistic structures with other similar or not languages. This way, the Albanian language can be analysed in typological and crosslingual settings. The values of the evaluation metrics for the artificial intelligence models in the experiments conducted in this research work increases at a rate similar to the ones observed in studies on other languages identified during the literature review, placing Albanian language at a comparable level. The results achieved highlight the impact of the architectures used and the different experimental configurations on the prediction of various linguistic labels. Finally, the integration of cognitive data to enhance NLP models open the way for future research on the Albanian language and the development of models of broader value for the interested research community.

Keywords: eye-tracking, mouse-tracking, natural language processing, part of speech tagging, dependency parsing, keyword extraction, Albanian language.

FALËNDERIME

Së pari dëshiroj të shpreh mirënjohjen time të thellë për udhëheqësen time shkencore, prof. dr. Elinda Meçe, e cila ka qenë një mbështetje e vazhdueshme gjatë gjithë viteve të studimeve të doktoratës, duke ofruar rregullisht ndihmën dhe këshillat e saj për çdo diskutim apo problematikë të shfaqur gjatë punës kërkimore. Si udhëheqëse shkencore, por edhe si Dekane e Fakultetit të Teknologjisë së Informacionit, ajo nuk ka kursyer asnjëherë mbështetjen për mbarëvajtjen e studimeve doktorale, si në planin profesional, ashtu edhe në atë njerëzor. Për mua, është një shembull frymëzues për t'u ndjekur, duke më udhëhequr dhe motivuar që në hapat e parë të karrierës sime akademike.

Një falënderim i veçantë i drejtohet Përgjegjësit të Departamentit të Inxhinierisë Informatike, prof. asoc. Hakik Paci, për përkrahjen dhe inkurajimin e vazhdueshëm për të çuar përpara dhe për të përfunduar çdo proces në kohën e duhur. Gjatë këtyre viteve kam pasur fatin të rrethohem nga kolegë të nderuar në Departamentin e Inxhinierisë Informatike, të cilët nuk kanë hezituar të ofrojnë mbështetje, këshilla dhe bashkëpunim të vazhdueshëm. E konsideroj veten me fat që çdo ditë kam mundësi të shkëmbej mendime dhe ide me ta, jo vetëm në lidhje me punën, por edhe përtej saj në diskutime akademike dhe profesionale. Falënderoj prof. dr. Aleksandër Xhuvani për fillimin e udhëheqjes së doktoratës dhe prof. asoc. Evis Trandafili për motivimin në fushën e gjurmimit të syrit.

Një falënderim i veçantë shkon për Zëvendësdekanen për Anën Shkencore në Fakultetin e Teknologjisë së Informacionit, dr. Nelda Kote, me të cilën kam pasur fatin të ndaj një interes të përbashkët kërkimor për përpunimin e gjuhës natyrore, veçanërisht për gjuhën shqipe. Bashkëpunimi me të, si edhe me koleget e gjuhësisë prof. asoc. Anila Çepani dhe prof. asoc. Adelina Çerpja, ka qenë një mundësi e çmuar për të zhvilluar më tej njohuritë e mia në këtë fushë ndërdisiplinore.

Frymëzimi për të punuar në një fushë kaq të veçantë si gjurmimi i syrit dhe ndërlidhja me përpunimin e gjuhëve natyrore është i lidhur ngushtë me pjesëmarrjen time në projektin COST MultiplEYE. Të qenët pjesë e këtij projekti më ka dhënë mundësi të njoh dhe të bashkëpunoj me studiues të shumtë nga kjo fushë. I jam mirënjohëse prof. dr. Lena Jäger dhe studenteve të saj të doktoratës për bashkëpunimin gjatë shkëmbimeve të mia studimore në Departamentin e Gjuhësisë Kompjuterike në Universitetin e Zyrihut.

Një falënderim i pamatshëm shkon për prindërit e mi. Çdo arritje akademike imja është edhe meritë e tyre dhe e lidhur ngushtë me përkrahjen që më kanë dhënë. Shpreh mirënjohjen time të sinqertë për familjen e bashkëshortit për inkurajimin dhe besimin gjatë këtij rrugëtimi akademik. Një falënderim për familjarët dhe miqtë të cilët kanë gjetur kohën për të lexuar dhe për të ofruar komentet dhe sugjerimet e tyre të vlefshme. Së fundmi, falënderimi im i veçantë i drejtohet Arditit për durimin, mirëkuptimin dhe mbështetjen e tij të vazhdueshme gjatë gjithë këtij rrugëtimi akademik.

Alba Haveriku

Tiranë, 2026

LISTA E FIGURAVE

LISTA E TABELAVE

LISTA E GRAFIKËVE

FJALOR TERMINOLOGJIK

TERMI ANGLISHT	TERMI SHQIP	KUPTIMI
Cognitive data	Të dhëna konjitive	Të dhëna që pasqyrojnë procese mendore si vëmendja, perceptimi dhe përpunimi i informacionit gjatë një aktiviteti, p.sh. leximi.
Cognitive Processes	Procese njohëse	Mekanizma mendorë që përfshijnë perceptimin, vëmendjen, kujtesën dhe kuptimin.
Data preprocessing	Parapërpunim i të dhënave	Procesi i pastrimit, filtrimit dhe strukturimit të të dhënave para analizës së mëtejshme.
Deep learning	Të mësuarit e thelluar	Nënfushë e të mësuarit të makinës e bazuar në rrjetet nervore me shumë shtresa.
Dwell time	Kohë qëndrimi	Koha totale e kaluar mbi një njësi tekstuale gjatë leximit.
Digital Object Identifier (DOI)	Identifikuesi i objektit digjital	Kod unik që përdoret për identifikimin dhe citimin e objekteve digjitale shkencore, si artikuj, kapituj librash ose të dhëna kërkimore.
Eye-tracking	Gjurmimi i syrit	Teknikë eksperimentale për regjistrimin dhe analizën e lëvizjeve të syrit dhe të vështrimit gjatë kryerjes së një detyre, si leximi i tekstit.
Feature extraction	Nxjerrje veçorish	Proces i të mësuarit të makinës ku të dhënat me shumë dimensione transformohen në një bashkësi më të vogël të dhënash informuese.
Fixation	Fiksim	Ndalesë e shkurtër e vështrimit mbi një fjalë ose zonë teksti.
Fixation duration	Kohëzgjatje e fiksimit	Koha gjatë së cilës vështrimi mbetet i palëvizshëm mbi një element të tekstit.
Focal points	Pika fokale	Zona ose pika ku përqendrohet vëmendja vizuale.
Gaze duration	Kohëzgjatje e vështrimit	Koha totale e kaluar mbi një fjalë ose segment teksti.

Hover time	Kohë qëndrimi e kursorit	Koha gjatë së cilës kursori i mausit qëndron mbi një zonë të caktuar të tekstit.
Human - Computer Interaction (HCI)	Ndërveprimi njeri-kompjuter	Fushë që studion ndërveprimin mes njerëzve dhe sistemeve kompjuterike.
Large Language Models (LLMs)	Modelet e Mëdha Gjuhësore (MMGJ)	Modele gjuhësore të avancuara të bazuara në sisteme inteligjente të të mësuarit të thelluar, të trajnuara mbi sasi shumë të mëdha teksti për të mësuar rregulla gjuhësore dhe për të gjeneruar ose analizuar tekst.
Low-resource languages	Gjuhë me burime të pakta	Gjuhë me mungesë korpusesh, mjetesh dhe të dhënash digjitale.
Machine learning	Të mësuarit e makinës	Nënfushë e inteligjencës artificiale e cila lejon sistemet të mësojnë nga të dhënat, të identifikojnë rregulla dhe të marrin vendime pa qenë të paraprogramuara.
Mouse tracking	Gjurmimi i kursorit të mausit	Teknikë për regjistrimin dhe analizën e lëvizjeve të kursorit në kohë gjatë ndërveprimit me tekstin ose ndërfaqen.
Mouse trajectory	Trajektorja e kursorit	Rruga e ndjekur nga kursori i mausit gjatë ndërveprimit me tekstin.
Multi layer neural network	Rrjet nervor me shumë shtresa	Model llogaritës i përbërë nga disa shtresa neuronesh artificiale, që përpunojnë informacionin në mënyrë hierarkike.
Named Entity Recognition (NER)	Njohja e entiteteve të emërtuara	Procesi i identifikimit dhe klasifikimit të njësive gjuhësore në një tekst që përfaqësojnë entitete të botës reale, si persona, vende, organizata, data dhe shprehje numerike, në bazë të kontekstit gjuhësor.
Natural Language Processing	Përpunimi i gjuhës natyrore (PGJN)	Fushë e inteligjencës artificiale që merret me analizën dhe përpunimin automatik të gjuhës njerëzore.
Open dataset	Bashkësi e hapur të dhënash	Të dhëna të publikuara për përdorim dhe verifikim të hapur kërkimor.

Part of speech tagging (PoS)	Etiketimi i pjesëve të ligjëratës	Procesi i identifikimit dhe caktimit të klasës leksiko-gramatikore të fjalëve në një tekst, në varësi të formës së tyre morfologjike dhe funksionit që kryejnë në kontekstin sintaksor.
Reading behaviour	Sjellja e leximit	Modeli i lëvizjeve të vështrimit ose të kursorit që tregon mënyrën se si një lexues përpunon tekstin.
Reading metrics	Metrika të leximit	Matje sasiore që përshkruajnë aspekte të leximit, si kohëzgjatja e fiksimeve ose kthimet prapa në tekst.
Reading task	Detyrë leximi	Aktivitet i strukturuar leximi i përdorur në eksperiment.
Regression (eye movement)	Regresion	Lëvizje e vështrimit drejt pjesëve të mëparshme të tekstit, që lidhet me ripërpunimin ose vështirësinë e kuptimit.
Saccade	Sakadë	Lëvizje e shpejtë e syrit midis dy fiksimeve gjatë leximit.
Scanning trajectory	Trajektorja e skanimit	Rruga e ndjekur nga vështrimi gjatë leximit të tekstit.
Self-paced reading	Lexim me ritëm të vetëkontrolluar	Metodë eksperimentale ku lexuesi kontrollon vetë ritmin e paraqitjes së tekstit, zakonisht duke kaluar nga një segment në tjetrin, ndërsa regjistrohet koha e leximit për secilin segment.
Spotlight	Zona e fokusit	Pjesë e kufizuar e tekstit ose e ekranit ku përqendrohet vëmendja në një moment të caktuar gjatë leximit.
Text segmentation	Segmentim i tekstit	Ndarja e tekstit në njësi më të vogla për analizë.
Time-series data	Të dhëna kohore	Të dhëna të organizuara si varg ngjarjesh të regjistruara në kohë.
Timestamp	Shënues kohor	Vlerë që shënon momentin e saktë kohor kur regjistrohet një ngjarje ose e dhënë.

KREU 1

HYRJE

1.1. Motivimi dhe objektivat e studimit

Gjatë dekadës së fundit, sistemet e përpunimit të gjuhës natyrore (PGJN) kanë shënuar një zhvillim të shpejtë, nxitur nga përparimet në algoritmet e të mësuarit të makinës (angl. *machine learning)* dhe nga rritja e kapaciteteve kompjuterike. Duke nisur nga detyra të thjeshta për përpunimin e tekstit dhe deri në procese më të ndërlikuara të kuptimit dhe gjenerimit të tij, PGJN përbën sot një komponent qendror në shumë aplikacione të inteligjencës artificiale.

PGJN është një fushë e Inteligjencës Artificiale (IA) që synon t'u mundësojë kompjuterëve përpunimin dhe kuptimin e gjuhës së shkruar dhe të folur (Khatter & Singh, 2022). Zhvillimi i kësaj fushe lidhet fillimisht me nevojën për komunikim më natyror ndërmjet përdoruesve dhe sistemeve kompjuterike, pa nevojën për të mësuar gjuhë specifike programimi. Kërkimi në fushën e PGJN-së ka karakter ndërdisiplinor, duke u zhvilluar në ndërthurjen e shkencave kompjuterike me gjuhësinë, psikologjinë, filozofinë dhe fusha të tjera të afërta. Studimet në këtë fushë realizohen nga profesionistë me formime të ndryshme dhe, përtej zhvillimit të zgjidhjeve teknologjike, kanë dhënë kontribute të rëndësishme edhe në thellimin e kuptimit të mënyrës se si truri i njeriut përpunon gjuhën.

Gjuha është një mjet i fuqishëm i komunikimit njerëzor në përditshmëri e cila mundëson shprehjen e koncepteve dhe dukurive komplekse në mënyrë të strukturuar. Megjithatë, larmia dhe ndryshueshmëria ndërgjuhësore e bëjnë modelimin e gjuhës të vështirë për modelet kompjuterike. Rregullat gjuhësore shoqërohen shpesh me raste përjashtimore, ambiguitete dhe variacione kontekstuale në të gjitha nivelet e përshkrimit gjuhësor, ndërsa interpretimi i nuancave semantike kërkon jo rrallë njohuri kulturore specifike. Në këtë aspekt, përpunimi i gjuhës dallon nga fusha të tjera të inteligjencës artificiale, si përpunimi i sinjaleve ose i imazheve, të cilat karakterizohen nga struktura më të qëndrueshme dhe rregullsi më të qarta formale.

Ekzistenca e mbi 7000 gjuhëve në botë e vështirëson ndërtimin e një modeli kompjuterik të vetëm që të jetë në gjendje të kuptojë dhe mësojë parimet bazë të përpunimit të gjuhës. Përgjithësimi i këtyre parimeve nga një gjuhë ose një grup gjuhësh çon shpesh në kufizime, pasi gjuhë të tjera paraqesin struktura të ndryshme gjuhësore, çka ndikon në uljen e performancës së modeleve në këto raste. Si rrjedhojë, PGJN-ja përbën një fushë në zhvillim të vazhdueshëm, e cila eksploron vazhdimisht të dhëna, metoda dhe modele të reja për të përballuar këtë larmi gjuhësore.

Modelet e mëdha gjuhësore (angl. *Large Language Models - LLMs*) kanë sjellë një revolucion në fushën e PGJN-së vitet e fundit. Ky zhvillim është nxitur kryesisht nga disponueshmëria e sasive shumë të mëdha të të dhënave tekstuale, si edhe nga rritja e ndjeshme e kapaciteteve kompjuterike. Arkitektura të tilla si GPT *(Generative Pre-Trained Transformers* – shqip *Transformer gjenerues i trajnuar paraprakisht)* dhe modeli LLaMA, kanë transformuar mënyrën se si trajtohet sot përpunimi i gjuhës, veçanërisht edhe në rastin e gjuhëve me burime të pakta. Një LLM është një model i inteligjencës artificiale i trajnuar në sasi të mëdha të dhënash tekstuale i aftë të kuptojë dhe të gjenerojë tekst në një mënyrë që i afrohet përdorimit njerëzor, duke u përdorur në detyra të ndryshme që lidhen me përpunimin e gjuhës.

Zhvillimi i këtyre modeleve duket sikur po ridrejton fushën e PGJN-së, nga modelet specifike (të lidhura me një gjuhë apo detyrë të caktuar), në modele të mëdha të cilat mund të përshtaten përmes teknikave si *prompting* (udhëzim i modelit përmes hyrjeve tekstuale) ose *fine-tuning* (ritrajnim i pjesshëm i modelit mbi të dhëna specifike) për të adresuar larmishmëri më të madhe detyrash. Megjithatë, krijimi i bashkësive të të dhënave dhe puna për gjuhët me burime të pakta mbeten ende thelbësore dhe të pazëvendësueshme. Modelet shumëgjuhëshe shpesh kanë performancë më të ulët për gjuhë si shqipja, duke qenë se janë trajnuar kryesisht në gjuhë me burime të pasura si anglishtja, gjermanishtja etj. Për rrjedhojë, ndërtimi i korpuseve dhe përshtatja e modeleve për gjuhët me burime të pakta mbetet domosdoshmëri për të kuptuar strukturën specifike të një gjuhe dhe për të arritur nivele të kënaqshme të performancës së modeleve në një gjuhë të caktuar.

Kërkimi në PGJN është i lidhur ngushtë me zhvillimet në fushën e të mësuarit të makinës, megjithatë njeriu mbetet në qendër të kësaj fushe. Gjuha krijohet, zhvillohet dhe përdoret nga njerëzit, ndërsa sistemet e PGJN-së ndërtohen për t'u përdorur drejtpërdrejt prej tyre. Njerëzit mësojnë, kuptojnë dhe prodhojnë kontekste gjuhësore në mënyrë të nënvetëdijshme. Edhe pse mekanizmat konjitivë që qëndrojnë në themel të këtyre proceseve nuk janë ende plotësisht të kuptueshëm, fakti është se mendja njerëzore vazhdon të tejkalojë performancën e modeleve gjuhësore në aftësitë e tyre përgjithësuese. Kjo ka sjellë që fusha si psikolinguistika dhe neurolinguistika të shërbejnë si burim frymëzimi për ndërtimin e modeleve kompjuterike më të avancuara (Beinborn & Hollenstein, 2024). Në mënyrë që modelet të arrijnë të kuptojnë gjuhën dhe të fitojnë aftësitë për t'u përshtatur në struktura dhe kontekste të reja, të dhënat konjitive, si ato të grumbulluara nga eksperimentet e gjurmimit të syrit dhe të kursorit të mausit, mund të përdoren për të afruar modelimin kompjuterik me proceset njerëzore të përpunimit të gjuhës dhe për të përmirësuar performancën e modeleve ekzistuese të PGJN-së.

Kërkimi në fushën e PGJN-së për gjuhën shqipe, si një gjuhë me burime të pakta, përballet me një sërë kufizimesh dhe problematikash. Janë bërë përpjekje të ndryshme për të studiuar aspekte të ndryshme të shqipes, por këto studime janë shpesh të kufizuara në shkallë, të mbyllura për përdorim të gjerë ose të pjesshme në aspektet gjuhësore që marrin

në konsideratë. Ndërkohë, eksperimentet psikolinguistike të lidhura me gjuhën shqipe mbeten të pakta dhe zakonisht të zhvilluara në forma minimale.

Fusha e gjurmimit të syrit dhe teknikave të ngjashme mbetet ende pak e zhvilluar për studimin e sjelljeve gjatë leximit në gjuhën shqipe. Në gjuhë me burime të shumta si anglishtja apo gjermanishtja përdorimi i të dhënave konjitive të mbledhura nga eksperimentet psikolinguistike ka treguar përmirësime të dukshme në performancën e modeleve ekzistuese të PGJN-së në nënfusha të ndryshme, përfshirë etiketimin e pjesëve të ligjëratës, përcaktimin e kategorive sintaksore, njohjen e entiteteve gjuhësore dhe zbulimin e gabimeve gramatikore. Rrjedhimisht, pritet të arrihen përmirësime të ngjashme edhe në kontekstin e gjuhës shqipe. Për më tepër, meqenëse të dhënat konjitive janë provuar se ndihmojnë në etiketimet gjuhësore (Barrett & Søgaard, 2015), ato mund të përdoren për të përshpejtuar procesin e etiketimit, duke sjellë përshpejtim në krijimin e korpuseve të etiketuara për gjuhën shqipe. Megjithatë, një nga pengesat kryesore në zhvillimin e eksperimenteve të gjurmimit të syrit është kostoja e lartë e pajisjeve të cilat shpesh nuk janë të aksesueshme për laboratorë të vegjël kërkimorë që merren me gjuhë me burime të pakta, si gjuha shqipe. Në mungesë të pajisjeve të gjurmimit të syrit me cilësi të lartë, alternativat me kosto të ulët paraqiten si zgjidhje të domosdoshme për zgjerimin e studimeve psikolinguistike dhe për rritjen e përfaqësimit të gjuhëve me pak burime në kërkimin shkencor.

Duke marrë në konsideratë kontekstin aktual të studimit të gjuhës shqipe lidhur me fushat e PGJN-së dhe eksperimenteve psikolinguistike për mbledhjen e të dhënave konjitive (si gjurmimi i syrit dhe i kursorit të mausit), ky punim synon arritjen e objektivave të mëposhtme:

1. Vlerësimin e situatës aktuale për zhvillimin e eksperimenteve psikolinguistike si gjurmimi i syrit dhe gjurmimi i mausit.
2. Vlerësimin e metodave për integrimin e të dhënave nga eksperimentet e gjurmimit të syrit dhe të mausit në modelet e të mësuarit të makinës.
3. Rishikimin sistematik të literaturës shkencore për integrimin e teknikave të gjurmimit të syrit në PGJN.
4. Renditjen e pajisjeve kryesore të përdorura për zhvillimin e eksperimenteve të gjurmimit të syrit dhe metodave alternative me kosto të ulët.
5. Përshtatjen e një metodologjie eksperimentale për të zhvilluar studime me kosto të ulët të përshtatshme për mbledhjen e të dhënave konjitive në gjuhë me burime të pakta.
6. Zhvillimin e eksperimenteve për gjurmimin e kursorit të mausit gjatë leximit të teksteve në gjuhën shqipe me pjesëmarrjen e të paktën 50 folësve amtarë.
7. Përcaktimin e një metodologjie për përpunimin dhe analizën e të dhënave të mbledhura nga eksperimentet e gjurmimit të mausit.

8. Zhvillimin e eksperimenteve për gjurmimin e syrit me pajisjen *EyeLink Portable Duo* gjatë leximit të teksteve në gjuhën shqipe me pjesëmarrjen e të paktën 20 folësve amtarë.
9. Përcaktimin e një metodologjie për përpunimin dhe analizën e të dhënave të mbledhura nga eksperimentet e gjurmimit të syrit.
10. Integrimin e të dhënave konjitive nga eksperimentet e gjurmimit të syrit dhe të mausit për përmirësimin e modeleve të PGJN-së, në detyra si etiketimi i pjesëve të ligjëratës, përcaktimi i kategorive sintaksore dhe identifikimi i fjalëve kyçe. Analiza e rezultateve të arritura nga konfigurimet e ndryshme eksperimentale.

1.2. Kontributet e studimit

Kontributi kryesor i këtij studimi qëndron në përcaktimin dhe vlerësimin e dy metodologjive eksperimentale, njëra me kosto të ulët dhe tjetra me kosto të lartë, për zhvillimin e eksperimenteve psikolinguistike për gjuhët me burime të pakta, si dhe në integrimin e të dhënave konjitive të përftuara nga këto eksperimente në modelet e PGJN-së. Përmbledhtas kontributet kryesore të punimit janë:

1. Hartimi i një metodologjie eksperimentale për eksperimentet psikolinguistike, e përshtatur për dy alternativa me kosto të ulët, si gjurmimi i kursorit të mausit dhe kosto më të lartë, si gjurmimi i syrit përmes pajisjes EyeLink Portable Duo.
2. Zhvillimi i eksperimenteve për gjurmimin e syrit dhe mausit me pjesëmarrës të cilët janë folës nativë të gjuhës shqipe.
3. Publikimi i hapur i bashkësisë së të dhënave të grumbulluara nga eksperimentet e gjurmimit të mausit dhe gjurmimit të syrit (përmes EyeLink Portable Duo) gjatë leximit të teksteve në gjuhën shqipe.
4. Krahasimi i metrikave të leximit të gjeneruara nga bashkësitë e të dhënave nga gjurmimi i syrit dhe i mausit, si edhe analizimi i lidhjes mes tyre.
5. Përcaktimi i një metodologjie të strukturuar për përpunimin dhe analizën e të dhënave konjitive**,** të mbledhura nga gjurmimi i mausit dhe i syrit.
6. Vlerësimi empirik i ndikimit të integrimit të metrikave të leximit në performancën e modeleve të PGJN-së, në tri detyra kryesore: etiketimi i pjesëve të ligjëratës, përcaktimi i kategorive sintaksore dhe identifikimi i fjalëve kyçe.

1.3. Struktura e disertacionit

Ky disertacion është i organizuar në shtatë kapituj të ndërtuar në mënyrë progresive nga kuadri teorik dhe metodologjik drejt analizës eksperimentale dhe integrimit të rezultateve në modelet e përpunimit të gjuhës natyrore (PGJN).

Kreu i parë shërben si hyrje e përgjithshme e disertacionit. Në këtë kapitull paraqitet konteksti i përgjithshëm i studimit, motivimi shkencor dhe praktik për zhvillimin e tij, si dhe problematika që lidhen me përpunimin e gjuhës shqipe në kuadër të PGJN-së.

Gjithashtu, përcaktohen qartë objektivat e studimit dhe kontributet kryesore të disertacionit, duke vendosur bazën konceptuale dhe metodologjike për kapitujt pasues. Në fund të kapitullit jepet një përmbledhje e strukturës së punimit.

Kreu i dytë trajton marrëdhënien ndërmjet proceseve konjitive njerëzore dhe përpunimit të gjuhës natyrore. Fillimisht prezantohet kuadri teorik i psikolinguistikës, procesi i kuptimit dhe gjenerimit të gjuhës dhe identifikimi i disa prej çrregullimeve që lidhen me përpunimin e gjuhës. Më pas përshkruhen teknikat kryesore eksperimentale të përdorura në studime psikolinguistike, duke vendosur theksin në eksperimentet e gjurmimit të syrit dhe të mausit të cilat do të trajtohen më në detaje në kapitujt në vijim. Kapitulli përfshin gjithashtu metoda alternative, si leximi me ritëm të vetëkontrolluar (angl. *self-paced reading)* dhe gjurmimi i kursorit të mausit, si dhe diskuton rolin e të dhënave konjitive në ndërtimin dhe përmirësimin e modeleve të të mësuarit të makinës.

Kreu i tretë paraqet një rishikim të detajuar të literaturës për të kuptuar situatën aktuale të punimeve në fushën e gjurmimit të syrit të lidhur me eksperimentet në lexim. Në këtë kapitull analizohen studime që demonstrojnë përmirësime në nënfusha të ndryshme të PGJN-së pas integrimit të të dhënave konjitive. Gjithashtu prezantohen korpuset kryesore shumëgjuhësore që përmbajnë të dhëna të gjurmimit të syrit gjatë leximit, si dhe diskutohen pajisjet kryesore të përdorura në këto studime. Një vëmendje e veçantë i kushtohet metodave dhe pajisjeve me kosto të ulët, të cilat janë veçanërisht të rëndësishme për studimin e gjuhëve me pak burime, si gjuha shqipe.

Kreu i katërt shpjegon në mënyrë të detajuar eksperimentin e zhvilluar për gjurmimin e mausit gjatë leximit. Gjurmimi i mausit përdoret si një alternativë me kosto të ulët në krahasim me gjurmimin e syrit. Në këtë pjesë të disertacionit paraqiten veçoritë teknike të eksperimentit për gjurmimin e mausit, bashkësia e të dhënave të mbledhura dhe analiza e sjelljes së pjesëmarrësve në eksperiment. Në këtë kapitull prezantohet tekstet dhe pyetjet kuptimore të përzgjedhura për t'u përfshirë në platformën eksperimentale. Më tej paraqitet metodologjia e ndjekur për parapërpunimin e të dhënave të grumbulluara. Korpusi i krijuar në këtë rast, publikohet dhe bëhet i disponueshëm në formën e tij bazë (koordinatat e grumbulluara), si edhe ofrohet mundësia e përdorimit të metodologjisë së prezantuar për të nxjerrë metrikat e leximit dhe përgatitjen e datasetit për integrimin në modele të tjera.

Kreu i pestë prezanton përdorimin e pajisjes EyeLink Portable Duo si një mundësi të mirë për zhvillimin e eksperimenteve me saktësi të lartë për gjurmimin e syrit. Përkatësisht, në këtë kapitull shpjegohet në detaje lloji i eksperimenteve që janë zhvilluar, qëllimi i tyre, dhe kriteret bazë që janë ndjekur. Gjithashtu prezantohen profili i pjesëmarrësve në eksperiment, shpërndarja në moshë, profili, kriteret e përfshirjes ose përjashtimore si edhe etika e ndjekur me të dhënat e mbledhura gjatë eksperimenteve. Ngjashëm me kapitullin e pestë, edhe në këtë kapitull paraqiten veçoritë teknike të platformës eksperimentale duke përdorur EyeLink Portable Duo, analiza e sjelljes së përdoruesve gjatë leximit, bashkësia e të dhënave të mbledhura dhe metodologjia e parapërpunimit të këtyre të dhënave.

Kreu i gjashtë ndahet në tre pjesë kryesore: 1. Një prezantim i kontekstit aktual të modeleve të PGJN për gjuhën shqipe, korpuseve të disponueshme dhe modeleve të përpunimit automatik të të dhënave; 2. Krahasimi i rezultateve të modeleve standarde për etiketimin e pjesëve të ligjëratës dhe kategorive sintaksore me ato të arritura nga përfshirja e metrikave të leximit (përfituar nga korpuset e paraqitura në kapitullin 4 dhe 5). 3. Krahasimi i rezultateve të modeleve standarde për identifikimin e fjalëve kyçe me ato të arritura nga përfshirja e metrikave të leximit (përfituar nga korpuset e paraqitura në kapitullin 4 dhe 5);

Kreu i shtatë prezanton rezultatet e arritura nga eksperimentet e zhvilluara, duke përfshirë analizën e të dhënave dhe krahasimin e performancës së modeleve të aplikuara. Në këtë kapitull diskutohen përmirësimet e mundshme që mund të ndërmerren për të optimizuar rezultatet, si edhe identifikohen kufizimet dhe sfidat e hasura gjatë eksperimenteve. Në bazë të kontributeve të sjella nga ky disertacion, vendosen disa pika të rëndësishme për t'u zhvilluar në të ardhmen. Synimi i këtij kapitulli është të ofrojë një pasqyrë të detajuar të arritjeve kryesore dhe të tregojë potencialin në zhvillime të reja dhe aplikime praktike në të ardhmen.

KREU 2

PROCESET KONJITIVE DHE PGJN

2.1. Hyrje

Objektivi kryesor i fushës së përpunimit të gjuhëve natyrore (PGJN), e njohur edhe si një nënfushë e gjuhësisë kompjuterike, është formalizimi i parimeve gjuhësore në mënyrë që ato të jenë të përfaqësueshme dhe të përpunueshme nga sistemet kompjuterike. Si fushë studimi, PGJN-ja ka marrë një rëndësi të madhe për shkak të nevojës për të grumbulluar dhe përpunuar sasi të mëdha informacioni në trajtë teksti të prodhuara nga faqet ueb, si edhe informacione në trajtë zëri apo video në sisteme të ndryshme komunikimi. Duke qenë se nxjerrja manuale e informacionit është proces i ngadalshëm, kërkon shumë kohë dhe shoqërohet me kosto të larta, lind nevoja për zhvillimin e sistemeve që mundësojnë përpunimin automatik të këtij informacioni. Përpunimi i gjuhës natyrore përfshin një sërë aktivitetesh analitike, të cilat zakonisht mund të organizohen në disa nivele: 1) *analiza leksikore* e cila përfshin analizën fonologjike dhe morfologjike; 2) *analiza morfosintaksore* e cila përfshin etiketimin e pjesëve të ligjëratës dhe kategorive sintaksore; 3) *analiza semantike* që studion kuptimin e fjalëve dhe 4) *analiza e diskursit* dhe pragmatizmi të cilat shqyrtojnë mënyrën se si kuptimi ndërtohet në kontekste të ndryshme komunikimi (Mishra & Bhattacharyya, 2018).

Modelet tradicionale të përpunimit të të dhënave gjuhësore janë zhvilluar kryesisht në bazë të rregullave gjuhësore (angl. *rule-based)* ose si sisteme statistikore, në të cilat modeli përpiqet të nxjerrë rregullsi gjuhësore nga një numër i madh shembujsh përmes modelimit matematikor. Performanca e këtyre qasjeve është përmirësuar ndjeshëm duke përdorur modele të bazuara në rrjete nervore artificiale, të cilat mundësojnë përfaqësime më të pasura dhe më fleksibile të gjuhës (Beinborn & Hollenstein, 2024). Sot, përdorimi i modeleve nervore të paratrajnuara të gjuhës është një nga standardet kryesore për detyra të ndryshme në fushën e përpunimit të gjuhëve natyrore, i mundësuar nga rritja e ndjeshme e madhësisë së korpuseve të trajnimit dhe nga përparimet në kapacitetet llogaritëse.

Në vitet e fundit, në fushën e PGJN-së është vënë në dukje potenciali i përfshirjes së të dhënave konjitive për përmirësimin e performancës së modeleve ekzistuese. Gjatë procesit të leximit, lëvizjet e syrit zhvillohen në mënyrë sekuenciale përgjatë tekstit, duke përfshirë fiksime mbi njësi të ndryshme leksikore. Studime të shumta psikolinguistike kanë treguar se metrikat e nxjerra nga lëvizjet e syrit përgjatë secilës fjalë, lidhen drejtpërdrejt me nivele të ndryshme të përpunimit gjuhësor, përfshirë strukturën sintaksore dhe organizimin semantik të tekstit (Barrett & Hollenstein, 2020). Të dhënat konjitive përbëjnë një bazë të

vlefshme për krahasimin midis mënyrës se si modelet kompjuterike përpunojnë tekstin dhe mënyrës se si informacioni i shkruar përpunohet nga njeriu. Këto të dhëna mund të mblidhen përmes metodave të ndryshme eksperimentale, si gjurmimi i syrit, gjurmimi i kursorit të mausit, matja e kohës së reagimit në eksperimentet e leximit me ritëm të vetëkontrolluar (angl. *self-paced reading*), si edhe përmes teknikave neurofiziologjike si EEG, MEG dhe fMRI. Duke qenë se modelet aktuale të të mësuarit të makinës nuk arrijnë ende të përfaqësojnë në mënyrë të plotë kompleksitetin e përpunimit njerëzor të gjuhës, integrimi i të dhënave që pasqyrojnë sjelljen gjatë leximit paraqet një qasje premtuese për rritjen e performancës së këtyre modeleve, si edhe për një kuptim më të thelluar të atyre aspekteve konjitive që aktualisht mungojnë në sistemet kompjuterike.

2.2. Gjurmimi i syrit

Gjurmimi i syrit është një metodë eksperimentale që mundëson regjistrimin e lëvizjeve të syve dhe të pozicionit të vështrimit gjatë kohës që zhvillojmë detyra të ndryshme konjitive, përfshirë leximin (Carter & Luke, 2020). Kjo metodë është një nga teknikat më të përdorura për të studiuar vëmendjen vizuale. Përdorimi i pajisjeve të specializuara për gjurmimin e syrit lejon analizën e leximit në dy dimensione kryesore: dimensionin kohor dhe atë hapësinor. Dimensioni kohor lidhet me kohën që një individ i caktuar shpenzon duke u përqendruar në një fjalë specifike gjatë leximit. Nga ana tjetër, dimensioni hapësinor lidhet me vendndodhjen specifike në ekran ku lexuesi përqendron vështrimin e tij.

Gjatë leximit, një lexues lëviz vështrimin nga një fjalë në tjetrën, duke u ndalur më shpesh te njësitë leksikore që mbartin informacionin kryesor të tekstit. Ky proces mund të përshkruhet përmes një sërë metrikash sasiore, të cilat nxirren nga të dhënat e mbledhura gjatë eksperimenteve të gjurmimit të syrit. Ndër metrikat kryesore përfshihen: kohëzgjatja e fiksimit (angl. *fixation duration*), e cila mat kohën e qëndrimit të vështrimit mbi një njësi të caktuar*; pikat e fokusit* (angl. *focal points*), që tregojnë zonat e përqendrimit vizual; sakadat (angl. *saccades*), të cilat përfaqësojnë lëvizjet e shpejta të syrit midis fiksimeve; trajektorja e skanimit (angl. *scanning trajectory*), që pasqyron rrugën e ndjekur nga vështrimi gjatë leximit; kohëzgjatja e vështrimit (angl. *gaze duration*), që përfshin kohën totale të fiksimeve mbi një fjalë para se vështrimi të largohet prej saj; si dhe koha e përgjithshme e vështrimit (angl. *overall viewing time*), e cila përmbledh të gjithë kohën e shpenzuar mbi një njësi tekstuale gjatë procesit të leximit.

Një *fiksim* (angl. *fixation)* është një periudhë kohe gjatë së cilës sytë qëndrojnë relativisht të palëvizshëm mbi një objekt ose zonë vizuale, duke mundësuar perceptimin dhe përpunimin e informacionit vizual (Rayner, 2009). Duke qenë se syri ka një hapësirë shikimi të vogël, sytë duhet të lëvizin në mënyrë të shpeshtë duke bërë që kohëzgjatja e fiksimeve të jetë relativisht e shkurtër. Kohëzgjatja e fiksimit mund të ndryshojë në varësi të faktorëve të ndryshëm, si tipi i stimulit vizual, qëllimi dhe kompleksiteti i detyrës që duhet të kryejë, si edhe aftësive dhe nivelit të vëmendjes së individit. Përafërsisht një kohë fiksimi llogaritet të zgjasë mes 180-330 milisekonda (Rayner, 2009).

Termi *sakadë* përdoret për të përshkruar lëvizjen e shpejtë të syrit nga një pikë fiksimi në tjetrën (Rayner, 2009). Gjatë një sakade syri është në lëvizje dhe nuk përpunon informacion vizual. Shpejtësia dhe kohëzgjatja e një sakade varet nga detyra që individi po kryen, p.sh. një sakadë gjatë leximit është relativisht e vogël në përmasë (rrotullim rreth 2 gradë) dhe zgjat afërsisht 30 milisekonda, ndërsa një sakadë në eksperimente për perceptimin e një imazhi është zakonisht më e madhe (rreth 5 gradë rrotullimi) dhe zgjat midis 40-50 milisekondave (Rayner, 2009).

Më poshtë është një shembull konkret dhe tipik për të ilustruar pikat e fiksimit dhe sakadat gjatë leximit të një fjalie në gjuhën shqipe:

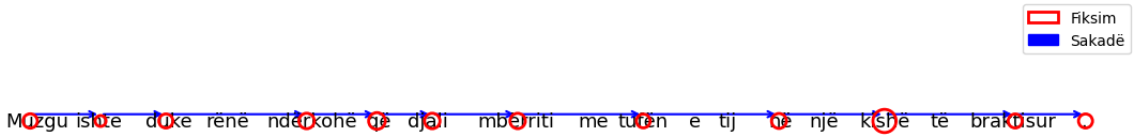


Figura 2.1. Shembull i fiksimeve dhe i sakadave në një fjali

Në shembullin e paraqitur në Figurën 2.1 vërejmë të shënuar në rrathë të kuq me madhësi të ndryshme fiksimet e para të pjesëmarrësit dhe në shigjetë blu drejtimin e lëvizjes së sakave. Vihet re se lexuesi kalon një pjesë të fjalëve pa u fiksuar gjatë procesit të leximit, duke u përqendruar më shumë në fjalët e kuptimit, ku në rastin e kësaj fjalie, fiksimi më i gjatë qëndron në fjalën "kishë".

2.2.1. Pajisjet e gjurmimit të syrit

Pjesa më e madhe e pajisjeve për gjurmimin e syrit janë të bazuara në kapjen e videove. Gjatë përdorimit të këtyre pajisjeve, një burim drite infra të kuqe, i padukshëm për syrin e njeriut, projektohet drejt syrit të pjesëmarrësit. Kjo dritë krijon reflektime karakteristike në kornenë e syrit dhe në bebe, duke bërë të mundur identifikimin e pozicionit të qendrës së bebes dhe të drejtimit të vështrimit. Duke analizuar këto reflektime në kohë reale, sistemi është në gjendje të llogarisë me saktësi pozicionin e pikës së vështrimit në ekran. Hapi i parë që kryhet gjatë zhvillimit të eksperimenteve mbi gjurmimin e syrit është kalibrimi, ku pjesëmarrësi në eksperiment udhëzohet të shikojë në disa pika vështrimi që shfaqen në mënyrë të njëpasnjëshme në qendrën e ekranit. Faza e kalibrimit ka mundësinë të përcaktojë sa e saktë është kapja e pikave të vështrimit. Përkatësisht, bëhet një përllogaritje se sa afër është pozicionin i pikës që shfaqet në ekran me vendin ku shikon individi.

Kalibrimi ndiqet nga faza e validimit, ku përcaktohet nëse kalibrimi është mjaftueshëm i saktë apo jo. Në shumicën e eksperimenteve përcaktohet një kufi minimal i saktësisë, ku nuk duhet të vijohet me eksperimentin nëse saktësia e kalibrimit bie poshtë këtij niveli minimal saktësie.

Ekzistojnë një larmi pajisjesh për gjurmimin e syrit, ku më të përdorshmet për eksperimentet në kushte laboratorike janë të prodhuara nga kompania Tobii (Tobii, 2025) dhe nga kompania SR Research (SR Research Ltd, 2016-2025). Përtej produkteve të këtyre

kompanive, kërkues nga laboratorë të ndryshëm janë gjithmonë në kërkim dhe zhvillim të pajisjeve me kosto më të ulëta apo softuer për të shndërruar kamerat web apo mobile në pajisje të mirëfillta për zhvillimin e eksperimenteve për gjurmimin e syrit. Një sërë kriteresh duhen marrë në konsideratë para se të zgjedhim pajisjen apo sistemin e duhur për zhvillimin e eksperimenteve. Në kreun 3.4 paraqiten më në detaje pajisjet më të përdorura nga kërkuesit e ndryshëm gjatë punimeve në këtë fushë.

Gjatë zhvillimit të eksperimenteve për gjurmimin e syrit për të siguruar që të dhënat e mbledhura reflektojnë në mënyrë të saktë proceset konjitive është e rëndësishme që të ndiqen në mënyrë rigoroze kritere për ruajtjen e cilësisë. Pika e parë kyçe që duhet përzgjedhur me kujdes është frekuenca e kampionimit. Frekuenca e marrjes së kampionëve nga pajisjet e gjurmimit të syrit nis nga 50 Hz deri në 2000 Hz. Kuptohet që sa më e lartë frekuenca e kampionimit aq më shumë informacion kemi dhe të dhënat e grumbulluara do të përfaqësojnë më mirë situatën eksperimentale.

Një pikë tjetër e rëndësishme është pjesa e stabilizimit të kokës. Disa nga pajisjet e gjurmimit të syrit vijnë së bashku me një vendqëndrim për të mbështetur kokën (angl. *head* dhe *chin rest*). Mbajtja e palëvizur e kokës gjatë eksperimenteve rrit nivelin e saktësisë së kapjes së lëvizjeve të syve. Në varësi të tipologjisë së eksperimenteve mundet ose jo të përfshihet edhe ky kriter si i domosdoshëm për të marrë të dhëna mjaftueshëm të sakta.

Gjatë grumbullimit të të dhënave të syrit mund të përzgjidhet si kriter ruajtja e të dhënave nga lëvizja e të dy syve ose vetëm njërit prej tyre. Në përgjithësi kjo nuk ka një ndikim shumë të madh meqë të dy sytë lëvizin njëkohësisht. Për të bërë më të menaxhueshme kapjen dhe ruajtjen e të dhënave, në eksperimente të gjata ndiqet praktika e ndjekjes së syrit dominant.

Së fundmi, disa kritere që shpesh mund të tejkalohen por janë të rëndësishëm në saktësinë e kapjes së bebes së syrit janë të ndërlidhur me nivelin e lodhjes së pjesëmarrësit, formën specifike të syrit, problematikave të syrit dhe përdorimin e produkteve të makijazhit. Për këtë arsye është e rëndësishme që të lajmërohen individët para se të paraqiten ditën e eksperimentit për të gjitha kriteret që duhet të kenë parasysh në varësi të eksperimentit.

2.2.2. EyeLink Portable Duo

Kjo pajisje ofron një sistem portabël për gjurmimin e syrit duke mundësuar një gjurmim të syrit në mënyrë të shpejtë dhe të saktë në një larmishmëri skenarësh kërkimorë. Ajo ofron një ndjekje monokulare ose binokulare të syve me një frekuencë kampionimi deri në 2000 Hz. Për të rritur saktësinë e pajisjes ofrohet mundësia e përdorimit të një stabilizuesi të kokës, në mënyrë që të minimizohen lëvizjet e pjesëmarrësit. Kjo pajisje ndriçon fytyrën e pjesëmarrësit me një dritë të padukshme infra të kuqe dhe një kamera e brendshme kap deri në 2000 imazhe në sekondë.

Imazhet e kamerës përpunohen nga kompjuteri bashkëngjitur me pajisjen i cilësuar si ‘Host PC’ për të identifikuar sytë dhe fytyrën e pjesëmarrësit, dhe duke lejuar matjen e

pikës ku po vështron lexuesi. Një kompjuter i dytë, i cilësuar si ‘Display PC’ ndërlidhet përmes lidhjes Ethernet me ‘Host PC’, duke mundësuar kalimin e informacionit kritik mes dy kompjuterëve. Kompjuteri ‘Display PC’ përdoret për të paraqitur stimujt e përzgjedhur tek pjesëmarrësit. Komunikimi i përhershëm midis ‘Display PC’ dhe ‘Host PC’, lejon kalimin në kohë reale të të dhënave dhe krijimin e logjikës së programimit për të dërguar mesazhe gjatë ngjarjeve (fiksimet, sakadat, paraqitjen e stimujve). I gjithë ky informacion ruhet i enkoduar në trajtë binare në file-n e fundit të krijuar nga eksperimenti për secilin pjesëmarrës. Konfigurimi fizik i nevojshëm për zhvillimin e eksperimenteve për gjurmimin e syrit përmes pajisjes EyeLink Portable Duo paraqitet në Figurën 2.2.

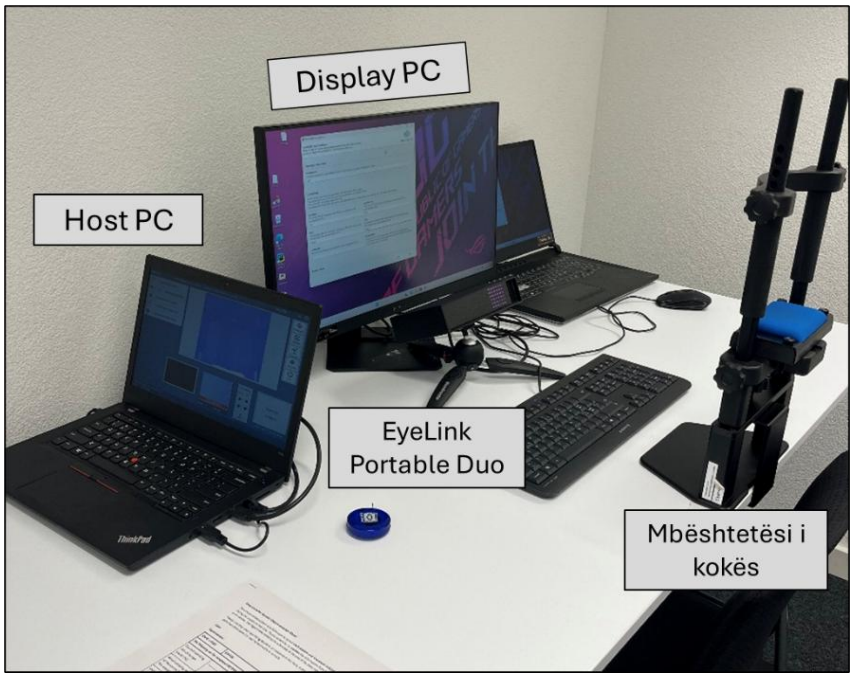


Figura 2.2. Paraqitja e konfigurimit fizik për eksperimentet me pajisjen EyeLink Portable Duo

Pajisja EyeLink Portable Duo ofron një sërë specifikimesh nga të cilat mund të përzgjidhen parametrat më të përshtatshëm për zhvillimin e eksperimenteve specifike. Në Tabelën 2.1 paraqiten specifikimet kryesore funksionale të pajisjes (SR Research Ltd, 2016-2025). Gjatë përdorimit të pajisjes përdoruesi ka akses në disa funksionalitete konfigurimi dhe kontrolli. Si fillim paraqitet ekrani i përzgjedhjes së konfigurimeve bazë ku përfshihen: përcaktimi i modalitetit të kapjes së syrit (monokular ose binokular), niveli i ndriçimit dhe mënyra e kalibrimit që do të ndiqet gjatë eksperimenteve, siç paraqitet në Figurën 2.3.

Gjithashtu në këtë fazë mund të verifikohet nëse pjesëmarrësi është pozicionuar saktë përpara pajisjes, dhe të rregullohet kapja e syrit/syve në mënyrë sa më të pastër. Nëse është

e nevojshme mund të rregullohen vlerat limit (angl. *threshold*) e kapjes së bebes dhe të reflektimit të kornesë (angl. *corneal reflection*), paraqitur në Figurën 2.4.

Formati fillestar i të dhënave të mbledhura nga gjurmimi i syrit përbëhet nga një seri kampionesh, ku çdo kampion përmban pikën e vështrimit të një ose të dy syve të paraqitur si koordinata x dhe y në pixel në ekranin e kompjuterit.

Tabela 2.1. Specifikime funksionale të pajisjes *EyeLink Portable Duo*

	Konfigurimi me mbështetës koke	**Konfigurimi i lirë**
Saktësia mesatare	0.25°- 0.5°	0.25°- 0.5°
Frekuenca e kampionimit	Monokular: 250, 500, 1000, 2000 Hz Binokular: 250, 500, 1000, 2000 Hz	Monokular: 250, 500, 1000 Hz Binokular: 250, 500, 1000 Hz
Parimi i kapjes së syrit	Reflektimi i kornesë së syrit	
Distanca mes kamerës dhe syrit	42-62 cm	
Përshtatshmëria me syzet	Shumë e mirë	
Lëvizshmëria e kokës	± 25 mm horizontalisht/vertikalisht	20cm x 20cm në distancën 52 cm

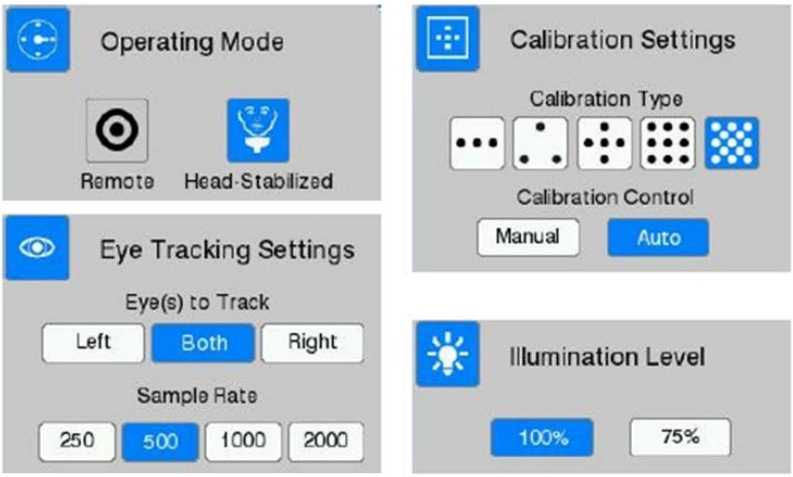


Figura 2.3. Ekranet e përzgjedhjes së konfigurimeve në pajisjen EyeLink Portable Duo

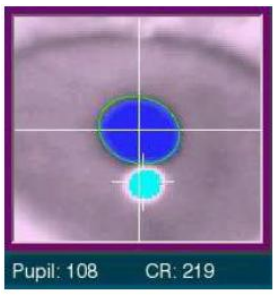


Figura 2.4. Rregullimi i vlerave të kapjes së bebes së syrit dhe reflektimit të kornesë së syrit në pajisjen EyeLink Portable Duo

Në varësi të tipit të pajisjes së përdorur mund të regjistrohen edhe të dhëna të tjera shtesë si p.sh. madhësia e bebes. Pajisja EyeLink Portable Duo i ruan të dhënat në një format `EDF`, i cili mund të transformohet në format `ASC` përmes mjetit `EDF2ASC` të ofruar bashkë me pajisjen. Skedarët EDF përmbajnë dy tipe të dhënash kryesore: kampionet e syrit (deri në 2000 kampione në sekondë) dhe ngjarjet (lëvizjet e syrit, fiksimet dhe sakadat). Secili kampion lidhet me kohën në milisekonda kur është kapur ky kampion, pozicionin e syrit në ato momente dhe madhësinë e bebes së syrit.

Gjatë përpunimit të këtyre të dhënave, hapi i parë është identifikimi i kampioneve që mund të grupohen në të njëjtin fiksim. Kjo ndodh kur këto kampione janë shumë afër në hapësirë me njëri-tjetrin. Ndërkohë kampionet të cilat janë afër në kohë por larg në hapësirë tregojnë se syri ka qenë në lëvizje me një nxitim të caktuar, duke i bërë këto kampione pjesë të një sakade. Disa përpunime bazike të të dhënave ofrohen nga pajisjet e gjurmimit të syrit, por në momentin që duam të nxjerrim metrika të caktuara është e nevojshme të përcaktohet një metodologji e qartë për përpunimin dhe analizën e këtyre të dhënave.

Në eksperimentet e gjurmimit të syrit gjatë leximit, është e rëndësishme të përcaktohen zonat e interesit (angl. *Areas of Interest* - *AOI*), të cilat zakonisht përkojnë me një karakter, një fjalë ose një segment të fjalisë. Kjo bën të mundur matjen e kohës gjatë së cilës një pjesëmarrës vështron një element të caktuar të tekstit, si edhe sa shpesh rikthehet tek ai element. Me përcaktimin e zonave të interesit mund të nxirren më pas metrika të ndryshme gjatë leximit si kohëzgjatja e leximit të parë (*first fixation duration*), numri i fiksimeve (*number of fixations*), koha totale e leximit (*total reading time*) etj.

2.3. Leximi me ritëm të vetëkontrolluar

Leximi me ritëm të vetëkontrolluar (angl. *self-paced reading* - *SPR*) është një teknikë eksperimentale që përdoret gjerësisht në psikolinguistikë për studimin e përpunimit të gjuhës gjatë leximit. Kjo metodë përqendrohet në analizën e leximit në njësi të vogla tekstuale, zakonisht fjalë ose segmente të shkurtra të fjalisë. Pjesëmarrësit në këtë tip eksperimenti duhet të shtypin një buton ose të klikojnë mausin për të shfaqur fjalën e radhës në ekran. Fjalët ose segmentet e tekstit janë fillimisht të fshehura dhe bëhen të dukshme

vetëm pas ndërveprimit të pjesëmarrësit me sistemin, zakonisht përmes shtypjes së një butoni. Në këtë mënyrë bëhet e mundur matja e kohës së reagimit, e cila interpretohet si koha që i nevojitet lexuesit për të përpunuar dhe kuptuar elementin e paraqitur të tekstit. Metoda e leximit me ritëm të vetëkontrolluar ka si avantazh kryesor faktin që është e lehtë për t'u zbatuar dhe mund të shpërndahet lehtësisht tek një numër i madh pjesëmarrësish, përfshirë edhe studime të realizuara në distancë.

Megjithatë kjo metodë ofron një saktësi më të ulët në krahasim me metodat e gjurmimit të syrit. Një nga kufizimet kryesore është se pjesëmarrësit nuk mund të rikthehen për të rilexuar segmentet e mëparshme të tekstit, çka e bën procesin e leximit më pak natyror. Gjithashtu, metoda nuk lejon të verifikohet nëse pjesëmarrësi e përdor kohën e regjistruar duke vështruar realisht elementin e shfaqur në ekran. Pavarësisht këtyre kufizimeve, kjo metodë është shumë e përdorur në zhvillimin e eksperimenteve psikolinguistike për shkak të thjeshtësisë dhe fleksibilitetit të saj.

Në këtë punim do të përqendrohemi në teknikat e gjurmimit të syrit dhe gjurmimit të mausit. Gjurmimi i mausit është përzgjedhur si alternativë me kosto të ulët meqenëse është provuar që ofron më shumë informacion se teknika e leximit me ritëm të vetëkontrolluar, duke sjellë të dhëna të përafërta në saktësi dhe në përfaqësim me ato të mbledhura gjatë eksperimenteve të gjurmimit të syrit (Wilcox, Ding, Sachan, & Jäger, 2024).

2.4. Gjurmimi i kursorit të mausit

Lëvizja e kursorit të mausit mundet të ofrojë informacion të rëndësishëm mbi mënyrën se si një person ndërvepron me një tekst të shkruar. Eksperimentet e gjurmimit të mausit janë alternativë e favorshme, sidomos e lidhur me zgjerimin e diversitetit të gjuhëve të studiuara dhe të dhënave psikolinguistike. Teoritë psikolinguistike kanë si qëllim të përshkruajnë gjuhën njerëzore në përgjithësi, dhe si rrjedhojë duhet të jenë të zbatueshme ndërgjuhësisht. Megjithatë, të dhënat aktualisht të disponueshme në këtë fushë janë të kufizuara si për nga shumëllojshmëria e gjuhëve, ashtu edhe për nga përfaqësimi i popullatave të studiuara. Gjurmimi i mausit për shkak se nuk kërkon infrastrukturë laboratorike të avancuar ofron një mjet të përshtatshëm për të zhvilluar studime ndërgjuhësore për përpunimin e fjalive, e cila si rrjedhojë do të ndihmonte në përgjithësimin e teorive psikolinguistike.

Studime të fundit vendosin theksin në përdorimin e gjurmimit të mausit si një metodë alternative e gjurmimit të syrit, pasi mund të kapë elemente të ngjashme të procesit të përpunimit të gjuhës gjatë leximit (Wilcox, Ding, Sachan, & Jäger, 2024). Ky studim është motivuar nga disponueshmëria e kufizuar e mjeteve për gjurmimin e syrit duke pasur si qëllim përfundimtar vërtetimin e potencialit të mjetit MoTR (Mouse Tracking for Reading/ Gjurmimi i Mausit për Lexim).

Gjurmimi i mausit është një teknikë ku gjatë zhvillimit të eksperimenteve pjesëmarrësit lëvizin kursorin e mausit për të ndjekur leximin e tekstit digjital. Kjo metodë paraqet disa avantazhe të rëndësishme metodologjike: nuk është invazive, vjen me kosto minimale dhe

është më e lehtë për t'u implementuar në krahasim me eksperimentet e gjurmimit të syrit. Megjithatë, një nga disavantazhet kryesore është fakti se kjo metodë nuk ofron të njëjtën saktësi me të dhënat nga eksperimentet e gjurmimit të syrit. Tri pikat kyçe në studimin e mjetit MoTR (Wilcox, Ding, Sachan, & Jäger, 2024) janë: saktësia, natyraliteti dhe kostoja e zbatimit. Mjeti MoTR është një alternativë më e lirë në krahasim me metodat e tjera ekzistuese për eksperimentimet e gjurmimit të syrit. Ky mjet ofron një paraqitje eksperimentale me një dizajn më natyral krahasuar me teknika të tjera të ngjashme si ato nëpërmjet përdorimit të kamerës web apo SPR (self-paced reading). Edhe pse nuk ofron të njëjta matje të sakta si në rastin e eksperimenteve për gjurmimin e syrit, kjo metodë është lehtësisht e aksesueshme dhe përfaqëson një alternativë të rëndësishme për studimin e gjuhëve me burime të kufizuara. Rezultatet e raportuara për përdorimin e mjetit MoTR në gjuhë të përhapura si anglishtja kanë treguar performancë të mirë, duke sugjeruar nevojën për vlerësimin e kësaj metode edhe në gjuhë të tjera.

Dy modele tipike të sjelljes së pjesëmarrësve gjatë këtyre eksperimenteve janë:

1. Krijimi i pikave të qëndrimit të kursorit dhe lëvizjeve të shpejta të mausit, të cilat janë analoge me fiksimet dhe sakadat e vëzhguara në gjurmimin e syrit;
2. Lëvizja më e ngadaltë e kursorit mbi tekst, me ndalesa më të gjata mbi fjalët që kërkojnë më shumë përpunim kognitiv.

Hapat e ndjekur në metodologjinë e përpunimit të të dhënave të mbledhura, të cilat kthejnë lëvizjet e syrit në kohë leximi për fjalë, përllogarisin sasinë e kohës gjatë së cilës kursori i mausit të pjesëmarrësit është më afër me një fjalë. Kjo qasje analitike lejon analizimin e të dy modeleve të leximit të përshkruara më sipër.

2.5. Metrikat e leximit në modelet e PGJN-së

Kërkimi në fushën e PGJN-së është i lidhur ngushtë me zhvillimet bashkëkohore në algoritmet e të mësuarit të makinë (angl. *machine learning*). Megjithatë, pavarësisht përparimeve teknologjike, njeriu mbetet elementi qendror i kësaj fushe. Gjuha natyrore gjenerohet dhe evoluon përmes ndërveprimit njerëzor, ndërsa mjetet për përpunimin e gjuhës përdoren drejtpërdrejt nga njerëzit. Si inputi ashtu edhe outputi i sistemeve të PGJN-së reflekton karakteristika të sjelljes dhe preferencave njerëzore. Për këtë arsye, përpunimi i gjuhës i varur nga përpunimi konjitiv i njeriut është shndërruar në një nga standardet në fushën e gjuhësisë kompjuterike.

Modelet kompjuterike të gjuhës përballen akoma me vështirësi në identifikimin dhe interpretimin e fenomeneve gjuhësore, të cilat përpunohen lehtë nga mekanizmat konjitiv të trurit të njeriut. Si rrjedhojë, psikolinguistika dhe neurolinguistika shikohen si frymëzim për ndërtimin e modeleve kompjuterike më të avancuara. Qasjet moderne synojnë të zhvillojnë arkitektura që nuk bazohen vetëm në modele statistikore të nxjerra nga bashkësi të kufizuara të të dhënave, por që gjithashtu marrin në konsideratë simulimin ose integrimin e sjelljes së lexuesve gjatë përpunimit të gjuhës.

Përdorimi i sinjaleve konjitive në modelet e PGJN-së paraqet një sërë sfidash metodologjike. Ndër problematikat kryesore mund të përmendim përmasat e kufizuara të bashkësive të të dhënave, disponueshmërisë e të dhënave dhe praninë e zhurmave në sinjalet e mbledhura gjatë eksperimenteve. Parapërpunimi i të dhënave konjitive përbën një hap të rëndësishëm dhe kompleks, duke përfshirë pika vendimmarrje si korrigjimi i zhurmave, agregimi mes pjesëmarrësve të ndryshëm, reduktimi i dimensioneve, ndërlidhja me stimujt etj.

Një nga qasjet e zakonshme në analizën e të dhënave konjitive është mesatarizimi i vlerave të sinjaleve të leximit midis pjesëmarrësve të ndryshëm. Kjo metodë është e përdorshme por shpesh herë mund të sjellë humbjen e karakteristikave individuale të pjesëmarrësve. Një alternativë tjetër është trajnimi i modelit mbi bashkësinë me të dhëna individuale për secilin pjesëmarrës. Kjo alternativë do të ishte e vlefshme në ndërtimin e modeleve më të gjeneralizueshme të cilët mund të kuptojnë të dhënat e rejat dhe të mundësojnë një vlerësim mes sjelljes së pjesëmarrësve të ndryshëm (Beinborn & Hollenstein, 2024).

Në punimin e tyre Brandl dhe Hollenstein (2022) analizojnë reflektimin e modeleve të leximit në grupe pjesëmarrësish me karakteristika të ndryshme dhe raportojnë ndryshime të konsiderueshme në sjelljen e leximit. Gjatë trajnimit të modeleve të mësuarit të makinës me sinjale të përpunimit konjitiv, rregullat e mësuara për një detyrë të caktuar, në mënyrë ideale, duhet të mund të zbatohen në një popullatë më të gjerë. Studimi dhe vlerësimi i skenarëve të ndryshëm duke veçuar një pjesëmarrës apo grupim pjesëmarrësish mund të ofrojë informacion të rëndësishëm në variueshmërinë e proceseve të leximit mes pjesëmarrësve. Zgjerimi i bashkësive të të dhënave konjitive po mundëson zhvillimin e eksperimenteve krahasuese në shkallë më të gjerë. Këto eksperimente përfshijnë krahasimin e përpunimit të gjuhës nga folës nativë ose jo nativë, krahasimin mes pjesëmarrësve me nivele të ndryshme arsimi, nivele të ndryshme lodhje etj.

Një sfidë tjetër gjatë punës me të dhënat konjitive është ndërlidhja mes stimujve eksperimental dhe tekstit të tokenizuar në një format që të jetë i mundur për tu kaluar si input në rrjetet nervore të përpunimit të gjuhës. P.sh. gjatë punës me të dhënat nga gjurmimi i syrit, duhet të vendoset si do të bëhet lidhja mes metrikave të fiksimit në nivel fjale në rastin e shprehjeve shumëfjalëshe.

Për më tepër, në rastin e eksperimenteve të gjurmimit të syrit apo mausit, shenjat e pikësimit nuk ndahen nga tokenat paraardhës ose pasardhës. Si pasojë, studiuesit duhet të vendosin nëse shenja e pikësimit do të trajtohen me vlerë zero të kohës së leximit, nëse koha e leximit do të ndahet mes fjalës dhe shenjës së pikësimit apo nëse do të duplikohet vlera e kohës së leximit. Përzgjedhja e secilës prej këtyre metodave duhet theksuar dhe arsyetuar duke qenë se ndikimi i tyre në vendimmarrjet e mëvonshme të modelit mbetet ende një fushë kërkimi aktive (Beinborn & Hollenstein, 2024).

Zhvillimet teknologjike në vitet e fundit tregojnë se pajisjet e gjurmimit të syrit po përmirësohen vazhdimisht duke e bërë më të lehtë mbledhjen e të dhënave konjitive. Kjo

prite të kontribuojë në krijimin e bashkësive më të mëdha të të dhënave duke rritur mundësinë për integrimin e metrikave të leximit në zhvillimin dhe vlerësimin e modeleve të PGJN-së. Për më tepër, përmirësimi i saktësisë dhe shpejtësisë së këtyre pajisjeve mund të lehtësojë zhvillimin e studimeve më të gjera mbi sjelljen gjatë leximit në popullata dhe gjuhë të ndryshme. Në këtë mënyrë, kombinimi i të dhënave konjitive me metodat e të mësuarit makinë mund të ndihmojë në ndërtimin e modeleve më të adaptueshme për analizën e gjuhës natyrore.

KREU 3

INTEGRIMI I TEKNIKAVE TË GJURMIMIT TË SYRIT NË PGJN

3.1. Hyrje

Koncepti i gjurmimit të syrit lidhet me detektimin dhe regjistrimin e lëvizjeve të syrit dhe shërben si dritare për studimin e vëmendjes vizuale dhe të proceseve njohëse të individit. Një numër i madh studimesh të zhvilluara kanë treguar se sjellja e syrit gjatë leximit pasqyron drejtpërdrejt proceset njohëse (angl. *cognitive processes)* të ndërtimit të kuptimit të tekstit, duke e bërë gjurmimin e syrit një metodë qendrore në kërkimin psikolinguistik.

Gjatë shqyrtimit të literaturës përkatëse është e rëndësishme të merret në konsideratë fakti që pajisjet për gjurmimin e syrit ndryshojnë ndjeshëm për sa i përket formës së instalimit, frekuencës së kampionimit, nivelit të saktësisë (angl. *accuracy*) dhe kostos. Pajisjet me kosto të lartë si p.sh. EyeLink 1000, përdoren kryesisht në kushte laboratorike dhe ofrojnë matje shumë të sakta dhe të qëndrueshme. Megjithatë, vitet e fundit pajisjet dhe sistemet alternative me kosto më të ulët kanë treguar që mund të arrijnë rezultate me një saktësi të pranueshme. Këto zhvillime kanë hapur rrugën për realizimin e studimeve më të aksesueshme, duke nxitur zgjerimin e kërkimit në gjurmimin e syrit edhe për gjuhë me burime të pakta, si gjuha shqipe.

Të dhënat e mbledhura gjatë eksperimenteve për gjurmimin e syrit janë të përdorshme në fusha të ndryshme si psikologji, gjuhësi, marketing, mjekësi dhe ndërveprimin njeri-kompjuter. Integrimi i këtyre të dhënave ka treguar një përmirësim të rezultateve të arritura nga modelet e përpunimit automatik të tekstit. Në literaturë janë raportuar një sërë studimesh të cilat hulumtojnë alternativat me kosto të ulët dhe mundësitë e studimeve ndërgjuhësore. Në fushën e përpunimit të gjuhës natyrale (PGJN) të dhënat e grumbulluara nga gjurmimi i syrit mund të përdoren për vlerësimin e rezultateve të modeleve të të mësuarit të makinës (angl. *machine learning*), për optimizimin e parametrave në modele të të mësuarit të thelluar (angl. *deep learning*), për studimin dhe përshtatjen e mekanizmave të vëmendjes në rrjetet nervore (angl. *neural networks*), si dhe për përmirësimin e detyrave konkrete gjuhësore, si etiketimi i pjesëve të ligjëratës (angl. *part-of-speech tagging*), njohja e entiteteve të emërtuara (angl. *named entity recognition*) dhe analiza e ndjenjave (angl. *sentiment analysis*).

Në vijim të këtij kapitulli paraqiten dhe diskutohen punimet kryesore të cilat dëshmojnë se integrimi i të dhënave të përftuara nga gjurmimit i syrit, si gjatë trajnimit ashtu edhe gjatë vlerësimit të modeleve, çon në përmirësime të ndjeshme të performancës në nënfusha të ndryshme të PGJN-së. Analiza e literaturës organizohet rreth katër këndvështrimeve themelore (Haveriku, Paci, Kote, Shasivari, & Meçe, 2025): 1) roli i gjurmimit të syrit në përmirësimin e nënfushave të PGJN-së; 2) korpuset shumëgjuhësore të gjurmimit të syrit

dhe rëndësia e tyre për analizat krahasuese; 3) pajisjet kryesore të përdorura për zhvillimin e eksperimenteve të gjurmimit të syrit; 4) përdorimi i pajisjeve me kosto të ulët për gjurmimin e syrit si alternativa të zbatueshme për kërkimin shkencor.

Në thelb të këtij studimi qëndrojnë katër pyetje kërkimore, të cilat kanë orientuar dhe motivuar punën e kryer:

- **Pyetje kërkimore 1 (PK1):** A sjell përdorimi i të dhënave nga gjurmimi i syrit përmirësim të performancës së modeleve të PGJN-së?
- **Pyetje kërkimore 2 (PK2):** A është i domosdoshëm krijimi i korpuseve shumëgjuhësore për krahasimin e sjelljes së leximit në gjuhë të ndryshme përmes gjurmimit të syrit?
- **Pyetje kërkimore 3 (PK3):** Cilat janë pajisjet më të përdorura dhe më të përshtatshme për zhvillimin e eksperimenteve për gjurmimin e syrit?
- **Pyetje kërkimore 4 (PK4):** A përbëjnë pajisjet dhe sistemet me kosto të ulët për gjurmimin e syrit një alternativë të vlefshme për lehtësimin zgjerimin e kërkimit eksperimental?

Studimet dhe publikimet e përmbledhura në këtë kapitull të punimit janë përzgjedhur nga baza të dhënash shkencore të njohura ndërkombëtarisht, si ScienceDirect, IEEE Xplore, ACM Digital Library, Springer, ACL Anthology, arXiv dhe Google Scholar.

Tabela 3.1. Kriteret përfshirëse dhe përjashtuese në rishikimin e literaturës

Baza digjitale e të dhënave	ScienceDirect, ACM Digital Library, IEEE Xplore, ACL Anthology, Springer, Google Scholar
Termat e kërkimit	eye-tracking, low-cost eye-tracking, natural language processing
Informacioni i ruajtur	Lista e autorëve, titulli i artikullit, viti i publikimit, identifikuesi i objektit digjital (angl. *digital object identifier (DOI))*, abstrakti.
Kriteret përfshirëse	▪ Studimi përshkruan përdorimin e teknikave për gjurmimin e syrit në nënfusha të PGJN. ▪ Studimi është publikuar në një revistë ose konferencë. ▪ Studimi është pjesë e një kërkimi primar.
Kriteret përjashtuese	▪ Studimi nuk lidhet me fushën e shkencave kompjuterike. ▪ Studimi nuk përshkruan metoda të ndërlidhura me PGJN ose gjurmimin e syrit. ▪ Studimi nuk është publikuar midis viteve 2015-2025.

Duke marrë në konsideratë zhvillimet e fundit teknologjike dhe metodologjike në këtë fushë, analiza është përqendruar në punime të botuara gjatë periudhës 2015-2025, me qëllim që të pasqyrohen prirjet më të fundit dhe përparimet kryesore në kërkimin mbi gjurmimin e syrit (Haveriku, Paci, Kote, & Meçe, 2024). Kriteret e përfshirjes dhe përjashtimit të studimeve të analizuara paraqiten në mënyrë të përmbledhur në Tabelën 3.1.

Numri fillestar i punimeve të mbledhura ka qenë 260 artikuj, nga të cilat pas ndjekjes së kritereve përfshirëse dhe përjashtuese, si edhe leximit të titullit dhe abstraktit u përzgjodhën 56 artikuj për analizim të mëtejshëm. Në Tabelën 3.2 është paraqitur numri i artikujve të përzgjedhur nga secila prej bazave digjitale të të dhënave.

Tabela 3.2. Përmbledhje e punimeve të përzgjedhura në rishikimin e literaturës

Baza digjitale e të dhënave	**Numri total i studimeve**	**Numri i studimeve të përzgjedhura**
Science Direct	150	25
ACM Digital Library	50	10
IEEE Xplore	35	7
Të tjera (ACL Anthology, Google Scholar, Springer)	25	14
Totali	**260**	**56**

3.2. Gjurmimi i syrit dhe përmirësimet në nënfushat e PGJN-së

Në këtë nënkapitull paraqiten dhe analizohen rezultatet e punimeve të cilat tregojnë se përdorimi i të dhënave nga gjurmimi i syrit arrin të përmirësojë ndjeshëm performancën e modeleve të PGJN-së, duke iu përgjigjur drejtpërdrejt pyetjes kërkimore 1 (PK1). Qëllimi është të evidentohet roli i sinjaleve konjitive si burim shtesë informacioni në detyra të ndryshme të përpunimit automatik të gjuhës.

Teknikat e gjurmimit të syrit janë përdorur në një sërë nënfushash të PGJN-së, përfshirë etiketimin e pjesëve të ligjëratës, analizën sintaksore, përmbledhjen automatike të tekstit, analizën e opinioneve dhe njohjen e entiteteve të emërtuara. Në vijim paraqitet një përmbledhje e studimeve përkatëse, ku për secilin punim specifikohen korpusi i përdorur, tipi i modelit të aplikuar, si edhe përmirësimi i arritur në performancë pas përfshirjes së të dhënave nga eksperimentet e gjurmimit të syrit. Një sintezë e këtij informacioni paraqitet në Tabelën 3.3.

Të dhënat e paraqitura në Tabelën 3.3. tregojnë se ndër vite një numër studimesh kanë shfrytëzuar të dhënat e mbledhura nga eksperimentet e gjurmimit të syrit për të pasuruar modelet e inteligjencës artificiale me informacion mbi sjelljen reale të leximit. Rezultatet

e raportuara në punime të ndryshme (Barrett & Søgaard, 2015), (Barret, Keller, & Sogaard, 2016), (Strzyz, Vilares, & Gómez-Rodríguez, 2019) dhe (Klerke & Plank, 2019) dëshmojnë se përfshirja e veçorive të vështrimit (angl. *gaze features*) mund të përmirësojë performancën e modeleve për etiketimin automatik të pjesëve të ligjëratës dhe lidhjeve sintaksore (angl. *dependency parser)*. Një aspekt veçanërisht i rëndësishëm i këtyre gjetjeve është fakti se përmirësimet vërehen edhe në skenarë me sasi të kufizuara të dhënash trajnuese, çka e bën përdorimin e të dhënave konjitive veçanërisht të vlefshëm për gjuhët me burime të pakta.

Një kontribut domethënës në këtë drejtim paraqitet nga Strzyz, Vilares dhe Gómez-Rodríguez (2019), të cilët demonstrojnë se veçoritë e vështrimit (angl. *gaze features)* mund të shfrytëzohen në mënyrë efektive për të përmirësuar analizën sintaksore, edhe në rastet kur të dhënat e gjurmimit të syrit janë të disponueshme vetëm gjatë fazës së trajnimit. Autorët propozojnë një qasje të bazuar në mësimin shumëdetyrësh (angl. *multi-task learning*), ku analiza e varësive sintaksore trajtohet si problem etiketimi sekuencial dhe sinjalet e gjurmimit të syrit mësohen si detyra ndihmëse. Në këtë mënyrë, informacioni konjitiv nuk përdoret drejtpërdrejt gjatë fazës së testimit, por ndikon në mënyrë indirekte në përditësimin e peshave të përbashkëta të rrjetit nervor gjatë trajnimit. Rezultatet eksperimentale, të vlerësuara si në të dhëna paralele ashtu edhe në të dhëna të ndara (angl. *disjoint datasets*), tregojnë përmirësime të qëndrueshme, megjithëse modeste, në saktësinë e analizës së lidhjeve sintaksore. Këto gjetje vërtetojnë se sinjalet e përftuara nga sjellja reale e leximit përmbajnë informacion të dobishëm për strukturën sintaksore dhe se integrimi i tyre në modelet e PGJN-së është i mundur edhe pa kërkesën për pajisje të gjurmimit të syrit në fazën e inferencës, çka e bën këtë qasje veçanërisht të përshtatshme për gjuhët me pak burime.

Mishra et al (2017), Barret et al (2018) si edhe Chen et al (2022) kanë prezantuar studime të cilat paraqesin avantazhet e përdorimit të dhënave nga gjurmimi i syrit në detyra të ndërlidhura me analizën e ndjenjës. Mishra et al (2017) në punimin e tyre kanë përdorur dy korpuse të disponueshme online për të detektuar sarkazmën dhe për të etiketuar ndjenjën e shprehur. Të dy korpuset janë zgjeruar duke shtuar si veçori të dhënat e vështrimit. Këto të dhëna janë kaluar në klasifikues si Naive Bayes, Support Vector Machine (SVM) dhe në një rrjet nervor me shumë shtresa (angl. *multi layer neural network)*. Rezultatet më të mira janë arritur nga përdorimi i klasifikuesit SVM i cili arriti një përmirësim performance me 6.1% në metrikën F1-score në rastin kur shtohen veçoritë e vështrimit. Gjithashtu autorët vendosin theksin në nevojën e përdorimit të pajisjeve mobile me kosto më të ulët në mënyrë që procesi i mbledhjes së të dhënave të vështrimit të kthehet në një proces me praktik dhe më të aksesueshëm.

Tabela 3.3. Renditja e punimeve kryesore të ndërlidhura me gjurmimin e syrit dhe nënfushat e PGJN-së

Autori dhe viti i publikimit	**Nënfusha e PGJN-së**	**Bashkësia e të dhënave**	**Tipi i modelit**	**Përmirësimi në saktësi**
(Barrett & Søgaard, 2015)	Identifikimi i kategorive sintaksore	250 fjali nga artikuj në revistën Wall Street, titujt kryesorë në revistën Wall Street, email, blogje në web dhe Twitter	- Supervised Logistic Regression	80.7-83.7%
(Barret, Keller, & Sogaard, 2016)	Etiketimi i pjesëve të ligjëratës	Dundee Treebank Penn Treebank	- Modele të të mësuarit të pakontrolluar - Të dhënat përdoren gjatë kohës së trajnimit dhe testimit	79.77-81.00%
(Mishra, Kanojia, Nagar, Dey, & Bhattacharyya, 2017)	Analiza e ndjenjave dhe detektimi i sarkazmës	**Korpusi i parë (**994 tekste: 383 pozitive, 611 negative, 350 sarkastike) **Korpusi i dytë (**1059 tekste: pozitive, negative, objektive)	- Support Vector Machine - Naïve Bayes - Multi-layers Neural Network	Një përmirësim me 3.7% në korpusin e parë dhe 9.3% në F1-score në korpusin e dytë.
(Rohanian, Taslimipoor, Yaneva, & Ha, 2017)	Parashikimi i shprehjeve me shumë fjalë (angl. *multi-word expressions)*	Korpusi Ghent Eye-Tracking (GECO)	- Conditional Random Fields (CRF) - Të dhënat përdoren gjatë kohës së trajnimit dhe testimit	F1-score ka vlerën 70.05% për fjalën e parë dhe 54.0% për fjalën e fundit në shprehje.

(Barret, Bingel, Hollenstein, Rei, & Søgaard, 2018)	Analiza e ndjenjave, detektimi i gabimeve gramatikore dhe detektimi i gjuhës abuzive	Korpusi i Dundee Korpusi i ZuCo	- Modeli BiLSTM	Zvogëlimi i gabimit mesatar (angl. *mean error)* nga modeli bazë me 4.5%.
(Strzyz, Vilares, & Gómez-Rodríguez, 2019)	Parashikimi i lidhjeve sintaksore	**Të dhëna paralele** Korpusi i Dundee **Disjoint data** Penn treebank, Dundee Treebank	- Modeli BiLSTM - Të dhënat përdoren gjatë kohës së trajnimit	Rasti i të dhënave në paralel: 85.36%- 85.61% Rasti i disjoint data: 93.98%-94.12%
(Klerke & Plank, 2019)	Etiketimi sintaksor	Korpusi i Dundee Korpusi Ghent Eye-Tracking (GECO)	- Modeli BiLSTM	Përmirësimi në rastin e korpusit Dundee: 95.25% - 95.48% Përmirësimi në rastin e korpusit GECO: 95.25% - 95.41%
(Hollenstein & Zhang, 2019)	Njohja e entiteteve të emërtuara (NER)	Korpuset: Dundee, GECO, ZuCo	- Modeli BiLSTM - Shtimi i veçorisë së vështrimit në një shtresë embedding	Metrika e saktësisë (angl. *precision*) 86.92% - 89.04%
(Chen, Mao, Liu, Zhang, & Ma, 2022)	Analiza e ndjenjës	Korpus i përbërë nga 1224 postime në gjuhën kineze me etiketat pozitive, neutrale, negative (408 për çdo kategori)	- Hierarchical Attention Networks - Reinforcement learning	Përmirësim me 2.13% të vlerës së F1-score
(Bautista & Naval, 2020)	Parashikimi i veçorive të vështrimit për secilën fjalë	Korpuset: ZuCo, Provo, UCL, GECO	- Modeli BiLSTM	Përmirësim me 2.67% në saktësi.

Në punimin e tyre, Barrett, Bingel, Hollenstein, Rei dhe Søgaard (2018) zhvillojnë eksperimente në tri nënfusha të PGJN-së: zbulimin e gabimeve gramatikore, zbulimin e gjuhës abuzive dhe analizën e ndjenjës. Një nga gjetjet kryesore të këtij studimi është demonstrimi i faktit se informacioni i nxjerrë nga vëmendja njerëzore (angl. *human attention)* e matur përmes të dhënave të gjurmimit të syrit gjatë leximit, mund të veprojë si një paragjykim induktiv (angl. *inductive bias)*, pra një supozim i integruar në model që orienton procesin e të mësuarit drejt strukturave më të përshtatshme. Ky paragjykim induktiv rezulton të jetë veçanërisht i dobishëm për përmirësimin e arkitekturave të rrjeteve nervore rikurrente (angl. *recurrent network architectures)*, modele nervore që përpunojnë sekuenca duke ruajtur informacion nga gjendjet e mëparshme, duke ndihmuar modelet të ndërtojnë përfaqësime më afër mënyrës se si njeriu përpunon tekstin gjatë leximit.

Në një linjë të ngjashme kërkimi, Chen, Mao, Liu, Zhang dhe Ma (2022) propozojnë një model që integron drejtpërdrejt informacion mbi sjelljen e leximit të individit (angl. *reading behavior*), duke kombinuar të dhëna konjitive me mekanizma të të mësuarit të thelluar (angl. *deep learning*). Rezultatet e raportuara tregojnë një përmirësim prej 2.13% në F1-score, duke konfirmuar se përfshirja e sinjaleve konjitive mund të rrisë ndjeshëm performancën e modeleve të PGJN-së.

Arritje të tjera të rëndësishme në këtë fushë janë prezantuar nga Rohanian, Taslimipoor, Yaneva dhe Ha (2017) të cilët përqendrohen në identifikimin automatik të njësive shumëfjalëshe (angl. *multiword expressions)*, pra shprehje të përbëra nga dy ose më shumë fjalë që funksionojnë si një njësi e vetme semantike ose sintaksore. Autorët analizojnë ndikimin e veçorive të vështrimit (angl. *gaze features)*, metrika të nxjerra nga sjellja e leximit si kohëzgjatja e fiksimit dhe rikthimet e vështrimit, në modelet ekzistuese për këtë detyrë dhe tregojnë se integrimi i këtyre veçorive përmirëson ndjeshëm aftësinë e modeleve për të dalluar kufijtë dhe strukturën e njësive shumëfjalëshe.

Nga ana tjetër, Hollenstein dhe Zhang (2019) fokusohen në përmirësimin e njohjes së entiteteve të emërtuara (angl. *named entity recognition - NER)*, duke shtuar të dhënat nga gjurmimi i syrit. Gjetjet kryesore të këtij studimi janë dyfish domethënëse: së pari, shtimi i veçorive të vështrimit në nivel fjale gjatë fazës së trajnimit e bën të panevojshme përfshirjen e këtyre veçorive në fazën e testimit, duke rritur praktikueshmërinë e modeleve; së dyti, modelet tradicionale të NER-it tejkalohen në performancë nga modelet e pasuruara me informacion konjitiv, çka dëshmon se sinjalet e vështrimit përmbajnë informacion të dobishëm për strukturën semantike të tekstit.

3.3. Korpuset shumëgjuhëshe mbi gjurmimin e syrit

Në këtë nënkapitull, paraqiten dhe analizohen punime të cilat theksojnë rëndësinë e krijimit të korpuseve shumëgjuhëshe studimin dhe krahasimin e të dhënave të gjurmimit të syrit gjatë leximit në gjuhë të ndryshme, duke iu përgjigjur drejtpërdrejt pyetjes kërkimore 2 (PK2).

Pjesa më e madhe e studimeve ekzistuese mbi gjurmimin e syrit është zhvilluar për gjuhë me përdorim të gjerë dhe burime të pasura si gjuha angleze dhe gjermane. Si

rrjedhojë, ka një rritje të nevojës për të zgjeruar kërkimin për gjuhët me përdorim më të vogël dhe me burime të pakta. Kjo krijon një nevojë të qartë për zgjerimin e kërkimit drejt gjuhëve të tjera, veçanërisht për të kuptuar nëse dhe në çfarë mase modelet e sjelljes gjatë leximit ndryshojnë në varësi të strukturës gjuhësore. Në studimet e fokusuara në lexim, qasjet ndërgjuhësore janë thelbësore për të analizuar variacionet që lindin nga dallimet tipologjike, ortografike dhe sintaksore mes gjuhëve të shkruara. Korpuset shumëgjuhëshe të gjurmimit të syrit mundësojnë jo vetëm krahasime sistematike të metrikave të leximit, por edhe testimin e përgjithësueshmërisë së modeleve konjitive dhe kompjuterike ndër gjuhë të ndryshme.

Në Tabelën 3.4 paraqiten disa nga korpuset ekzistuese të gjurmimit të syrit, njëgjuhëshe dhe shumëgjuhëshe, të cilat janë përdorur në studime psikolinguistike dhe në integrimin e të dhënave konjitive në modelet e PGJN-së.

Raymond et al. (2023) kanë prezantuar një korpus me burim të hapur i cili përmban të dhëna nga leximi, ku përfshihen disa gjuhë pak të studiuara në fushën e gjurmimit të syrit. Autorët identifikojnë dy arsye kryesore për diversitetin e limituar në studimet për gjurmimin e syrit: 1) kufizimet ekonomike (të ndërlidhura me kostot e larta të pajisjeve për gjurmimin e syrit); 2) shumica e pajisjeve për gjurmimin e syrit janë të përqendruara në studimin e gjuhëve latine. Puna në të ardhmen sipas tyre do të përfshijë të dhëna për gjuhë pak të studiuara si gjuha ukrainase, telugu dhe gjuhët indigjene afrikane.

Hollenstein et al. (2018) kanë krijuar një korpus të quajtur ZuCo (Zurich Cognitive Language Processing Corpus), ku kombinojnë informacionin e marrë nga gjurmimi i syrit me atë të mbledhur nga një pajisje EEG (Electroencephalography – matja e aktivitetit elektrik të trurit). I gjithë korpusi është i përbërë nga të dhënat e gjurmimit të syrit dhe EEG për rreth 21.629 fjalë me 154.173 fiksime të mbledhura në total. Puna e tyre duke integruar të dhënat e vështrimit dhe EEG, është një vlerë e shtuar për punimet në të ardhmen dhe përmirësimin e modeleve të mësimit të makinës për detyra si analiza e ndjenjave, njohja e entiteteve të emërtuara dhe zbulimi i lidhjeve semantike mes entiteteve.

Përveç korpuseve të lartpërmendura, një tjetër korpus i rëndësishëm është dhe RatrOS i zhvilluar nga Leal et al. (2022), i cili përmban të dhëna nga eksperimentet e gjurmimit të syrit nga studentë që lexojnë tekste në portugalishten braziliane. Ndërkohë, Slegelman et al. (2022) kanë analizuar 13 gjuhë të ndryshme, duke eksploruar ndryshimet e sjelljes gjatë leximit në secilën prej këtyre gjuhëve. Ata kanë zhvilluar MECO (*Multilingual Eye-tracking Corpus*), një korpus shumëgjuhësh me të dhëna nga leximi i syrit për të diferencuar mënyrat e ndryshme të sjelljes gjatë leximit mes gjuhëve të përfshira në eksperimente. Një nga rezultatet kyçe të kërkimit të tyre ishte fakti se norma e kalimit të fjalëve (angl. *skipping rate)* midis gjuhëve të ndryshme ka një diferencë të konsiderueshme, një fakt i cili lidhet me shpërndarjen e gjatësisë së fjalëve ndër gjuhët e ndryshme të përfshira në korpus. Autorët sugjerojnë si mundësi për punë në të ardhmen studimin e ndikimit të frekuencës dhe të gjatësisë së fjalëve, duke vendosur theksin në probabilitetin e kalimit të një fjale dhe probabilitetin e fiksimit, si dy veçori të cilat mund të ndryshojnë në kontekste të ndryshme gjuhësore.

Tabela 3.4. Korpuset një dhe shumëgjuhëshe për gjurmimin e syrit

Korpusi	**Numri i pjesëmarrësve në eksperiment**	**Gjuhët e përfshira**
(Raymond, Moldagali, & Madi, 2023)	40 pjesëmarrës 97 eksperimente	anglisht, spanjisht, kinezisht, gjuha hindi, rusisht, arabisht, japonisht, gjuha kazakh, urdu, gjuha vietnameze
Dundee (Kennedy, Hill, & Pynte, 2003)	10 pjesëmarrës	anglisht, frëngjisht
GECO (Cop, Dirix, Drieghe, & Duyck, 2017)	33 pjesëmarrës	anglisht, gjuha holandeze
ZuCo (Hollenstein, et al., 2018)	12 pjesëmarrës	anglisht
MECO (Slegelman, et al., 2022)	580 pjesëmarrës	gjuha holandeze, anglisht, gjuha estoneze, finlandisht, gjermanisht, greqisht, hebraishtja, italisht, gjuha koreane, norvegjisht, spanjisht, turqisht
Korpusi i fjalive në rusisht (Laurinavichyute, Sekerina, Alexeeva, Bagfasaryan, & Kliegl, 2018)	96 pjesëmarrës	rusisht
RatrOS (Leal, Lukasova, Carthery-Goulart, & Aluisio, 2022)	393 pjesëmarrës	portugalishte braziliane
Provo (Luke & Christianson, 2017)	84 pjesëmarrës	anglisht
WebQAmGaze (Ribeiro, Brandl, Søgaard, & Hollenstein, 2023)	332 pjesëmarrës	anglisht, spanjisht dhe gjermanisht
CopCo (Hollenstein, Barrett, & Bjornsdottir, 2022)	22 pjesëmarrës	gjuha daneze
RaCCooNS (Frank & Aumeistere, 2023)	37 pjesëmarrës	gjuha holandeze

Studime të tjera prezantojnë rezultatet e krahasimit midis gjuhëve të ndryshme. Barret et al. (2016) kanë studiuar korrelacionin midis pjesëve të ligjëratës dhe të dhënave të vështrimit në gjuhën angleze dhe frënge. Autorët tregojnë se të dhënat e vështrimit për gjuhën angleze mund të ndihmojnë në përmirësimin e modeleve të PGJN-së të krijuar për analizën e gjuhës frënge, duke treguar se modelet e etiketimit të pjesëve të ligjëratës me të dhëna për vështrimin të përfshira në to, mund të aplikohen në gjuhë të ndryshme. Eksperimentet e zhvilluara në këtë studim përdorin të dhëna nga korpusi shumëgjuhësh Dundee (Kennedy, Hill, & Pynte, 2003).

Hollestein et al. (2021) në punimin e tyre kanë analizuar performancën e disa modeleve shumëgjuhësore në parashikimin e sjelljes gjatë leximit. Gjuhët e përfshira në këtë studim janë gjuha holandeze, gjermane, angleze dhe ruse. Autorët përdorin modelin BERT (*Bidirectional Encoder Representations from Transformers)* dhe XLM (*Cross-lingual Language Model – Model gjuhësor shumëgjuhësh*) për të parashikuar veçoritë e vështrimit dhe për të provuar se modeli BERT ofron një përgjithësim më të mirë mes gjuhëve të ndryshme. Modeli XLM arrin një nivel më të mirë performance në rastet kur të dhënat në dispozicion janë në sasi më të vogël. Rezultatet e arritura nga autorët e këtij studimi tregojnë se modelet shumëgjuhësore mund të parashikojnë me sukses pikat e vështrimit gjatë leximit edhe në rastet kur gjuhët e përdorura nuk janë përfshirë në fazën e *fine-tuning* (përshtatje e specializuar e modelit*).* Edhe pse studimi është i limituar në katër gjuhë kryesore, ai tregon se është e mundur të përgjithësohen rezultatet nga një gjuhë në tjetrën.

Rezultatet e arritura nga analiza e këtyre studimeve sugjerojnë se ekziston një veçori universale në proceset konjitive të lidhura me kohën e leximit, pavarësisht diversitetit gjuhësor. Megjithatë, autorët theksojnë nevojën për studime më të zgjeruara për të përfshirë më shumë gjuhë dhe karakteristika të strukturave gjuhësore të cilat mund të ndikojnë në mënyrën e leximit. Një nga avantazhet kryesore të korpuseve shumëgjuhëshe të gjurmimit të syrit është mundësia për të krahasuar drejtpërdrejt lëvizjet e vështrimit në gjuhë të ndryshme, duke identifikuar faktorë specifikë për secilën gjuhë. Studime të mëparshme kanë provuar se disa metrika, si kohëzgjatja e fiksimeve apo gjatësia e sakadave janë të njëtrajtshme në disa gjuhë të ndryshme (Slegelman, et al., 2022). Megjithatë, aspekte të tjera të leximit të cilat mund të studiohen më tej përfshijnë veçoritë e ndërlidhura me ortografinë, fonologjinë dhe sintaksën.

Në studime të zhvilluara nga (Barrett, Bingel, Keller, & Søgaard, 2016) dhe (Hollenstein, Pirovano, Zhang, Beinborn, & Jäger, 2021) është provuar se veçoritë e marra nga të dhënat e gjurmimit të syrit mund të përgjithësohen në disa gjuhë duke lehtësuar parashikimin e mënyrës së leximit në kontekste shumëgjuhësore. Studimi shumëgjuhësor i të dhënave të ndërlidhura me gjurmimin e syrit është një fushë në zhvillim e sipër, duke sjellë domosdoshmërinë e zgjerimit të korpuseve shumëgjuhësore për të përmirësuar modelet kompjuterike si sistemet e përkthimit të automatizuar apo aplikimet e strategjive për të mësuar gjuhët e reja.

3.4. Pajisjet kryesore për gjurmimin e syrit

Në këtë pjesë paraqesim një përmbledhje të pajisjeve kryesore të cilat përdoren për zhvillimin e eksperimenteve të gjurmimit të syrit, me synimin për t'iu përgjigjur pyetjes kërkimore 3 (PK3).

Pas analizës së disa nga punimeve kryesore është e rëndësishme të paraqesim dhe një përmbledhje të pajisjeve më të përdorura për të krijuar korpuset me të dhënat e mbledhura gjatë eksperimenteve të gjurmimit të syrit.

Të dhënat e përmbledhura dhe të paraqitura në Tabelën 3.5 tregojnë se pajisja *EyeLink 1000+* është më e përdorura në studimet ekzistuese, kryesisht për shkak të saktësisë dhe frekuencës së lartë të kampionimit që ofron, duke e bërë veçanërisht të përshtatshme për studime laboratorike me kërkesa të larta metodologjike. Pas saj, një përdorim të gjerë gjejnë edhe pajisjet e prodhuesit *Tobii*, të cilat ofrojnë një kompromis midis saktësisë, fleksibilitetit eksperimental dhe kostos.

Tabela 3.5. Pajisjet kryesore për gjurmimin e syrit

Emri i pajisjes	**Prodhuesi**	**Frekuenca e kampionimit (Hz)**	**Saktësia**	**Punimet e ndërlidhura**
Tobii X120	Tobii	60/ 120	0.5 gradë	(Barrett & Søgaard, 2015), (Bautista & Naval, 2020)
Tobii T120	Tobii	120	0.5 gradë	(Mishra, Kanojia, Nagar, Dey, & Bhattacharyya, 2017)
Tobii X2-30 *(Zëvendësuar nga Tobii Pro Nano)*	Tobii	30	0.5 gradë	(Chen, Mao, Liu, Zhang, & Ma, 2022)
Tobii Pro Nano	Tobii	60	0.3 gradë në kushte optimale	-
Tobii Pro Spark	Tobii	60	0.4 gradë në kushte optimale	-
Tobii Pro Fusion	Tobii	30/ 60/ 120/ 250	0.3 gradë në kushte optimale	-
Tobii Pro Spectrum	Tobii	1200	0.3 gradë në kushte optimale	-

EyeLink II	SR Research	500	0.50 gradë	-
EyeLink Portable Duo	SR Research	2000	0.25-0.50 gradë	(Slegelman, et al., 2022)
EyeLink 1000	SR Research	250/500/1000 /2000	0.25 to 0.5 gradë	(Mishra, Kanojia, Nagar, Dey, & Bhattacharyya, 2017), (Cop, Dirix, Drieghe, & Duyck, 2017), (Slegelman, et al., 2022), (Leal, Lukasova, Carthery-Goulart, & Aluisio, 2022)
EyeLink 1000+	SR Research	2000	0.25-0.50 gradë	(Raymond, Moldagali, & Madi, 2023), (Hollenstein, et al., 2018), (Slegelman, et al., 2022), (Laurinavichyute, Sekerina, Alexeeva, Bagfasaryan, & Kliegl, 2018), (Luke & Christianson, 2017) (Hollenstein, Barrett, & Bjornsdottir, 2022), (Frank & Aumeistere, 2023)
Pupil Neon	Pupil Labs	220	2.4 gradë e pakalibruar	-
Pupil Core	Pupil Labs	200	0.6 gradë	-
Logic One	EyeLogic	60/120/250	<0.5 gradë	-
Lite	EyeLogic	60/120	<0.5 gradë	-

3.5. Përdorimi i pajisjeve me kosto të ulët për gjurmimin e syrit

Në këtë pjesë paraqesim rezultatet e punimeve të cilat vërtetojnë se pajisjet dhe sistemet për gjurmimin e syrit me kosto të ulët janë një alternativë e mirë për lehtësimin e zhvillimit të eksperimenteve, duke iu përgjigjur pyetjes kërkimore 4 (PK4).

Përdorimi i të dhënave të mbledhura nga eksperimentet e gjurmimit të syrit ka treguar se sjell përmirësime të ndjeshme në algoritmet e përpunimit automatik të tekstit. Pajisjet e paraqitura në Tabelën 3.5 janë pjesë kyçe e kërkimit në fushën e gjurmimit të

syrit. Edhe pse mënyra më e mirë për të mbledhur të dhëna me saktësi të pranueshme është përdorimi i pajisjeve me frekuencë kampionimi të lartë shumë studime janë kryer për të vlerësuar mundësinë e përdorimit të kamerave të thjeshta apo ato të përfshira në pajisjet telefonike për të zhvilluar eksperimente për gjurmimin e syrit.

Vështirësitë praktike dhe kostot e larta për zhvillimin e eksperimenteve për gjurmimin e syrit po nxisin gjithnjë e më shumë kërkimin drejt zhvillimit të sistemeve me kosto të ulët. Në Tabelën 3.6 listohen disa nga korpuset kryesore të grumbulluara duke përdorur kamerën ueb ose mobile.

Një nga kontributet më të hershme dhe më me ndikim në këtë drejtim është puna e Krafka et al. (2016) të cilët krijuan korpusin e parë me përmasë të madhe të grumbulluar nga eksperimentet e zhvilluara në pajisje mobile. Bashkësia e tyre e të dhënave, e njohur ndryshe me emrin GazeCapture, përmban të dhëna të mbledhura nga 1450 pjesëmarrës dhe rreth 2.5 milion korniza video (angl. *frames*). Autorët propozuan modelin iTracker, të bazuar në rrjete nervore konvolucionale (angl. *Convolutional Neural Networks-CNN*), i cili arriti një gabim mesatar parashikimi prej 1.71cm në pajisjet mobile dhe 1.34cm gjatë fazës së kalibrimit (angl. *calibration*). Për më tepër, veçoritë e mësuara nga ky model u testuan për aftësinë e tyre për përgjithësim në bashkësi të tjera të dhënash.

Në një punim më të vonshëm, Gunawardena et al. (2022) propozuan një arkitekturë mobile edge nëpërmjet të cilës kanë arritur të përafrojnë pikat e vështrimit në pajisjet mobile. Autorët kanë përdorur 1.4 milion korniza, pra imazhe të njëpasnjëshme të kapura nga videoja për detektimin e fytyrës dhe syrit nga korpusi i GazeCapture (Krafka, et al., 2016) dhe kanë krahasuar rezultatet nga përdorimi i katër arkitekturave CNN: AlexNet (Krizhevsky, Sutskever, & Hinton, 2017), LeNet-5 (Lecun, Bottou, Bengio, & Haffner, 1998), MobileNet (Howard, et al., 2017), dhe ShuffleNet (Zhang, Wang, & Liu, 2018). Përveç krahasimit të disa algoritmeve, autorët kanë testuar performancën e modelit të tyre në dy mënyra: në vetë pajisjen mobile dhe duke përdorur edge computing (përpunim i të dhënave jashtë pajisjes, por pranë burimit). Arritjet më kryesore të punës së tyre janë: a) Modelet MobileNet-V3 dhe ShuffleNet-V2 mundësojnë një gabim më të vogël në parashikim; b) Edge computing përmirëson ndjeshëm përvojën e përdoruesit në aplikacionet mobile për gjurmimin e syrit.

Një kontribut i veçantë në kontekstin e leximit është paraqitur nga Ribeiro et al. (2023) të cilët kanë krijuar një korpus shumëgjuhësor të cilësuar me emrin WebQAmGaze, një korpus i përbërë nga të dhëna të grumbulluara nga kamera të thjeshta ueb. Korpusi përmban të dhëna nga leximi i natyrshëm në anglisht, spanjisht dhe gjermanisht. Kërkuesit kanë krahasuar të dhënat e mbledhura nga kamera ueb me të dhënat e mbledhura nga pajisjet për gjurmimin e syrit me kosto të lartë, në mënyrë që të vlerësojnë saktësinë e kësaj qasjeje alternative.

Ndryshe nga punimet e tjera të përmendura më lart, të cilat fokusoheshin kryesisht në vështrimin e imazheve ose videove, WebQAmGaze është punimi i parë kërkimor i cili grumbullon të dhënat nga vështrimi në nivel fjale duke përdorur kamerën ueb. Edhe pse rezultatet e raportuara janë inkurajuese, (Ribeiro, Brandl, Søgaard, & Hollenstein, 2023) theksojnë se puna e ardhshme duhet të përqendrohet në zgjerimin e korpusit, përmirësimin e algoritmeve për parashikimin e pikave të vështrimit dhe integrimin e

modeleve të avancuara gjuhësore, si BERT**,** për të rritur më tej saktësinë dhe përdorshmërinë e të dhënave.

Tabela 3.6. Bashkësitë e të dhënave nga eksperimentet e gjurmimit të syrit përmes kamerës ueb apo mobile

Emri	**Nr. pjesëmarrës/ Madhësia e datasetit**	**Gabimi në parashikim**	**Karakteristika**
GazePoint (Krafka, et al., 2016)	1.450 pjesëmarrës 2.5 milion frames	Gabimi në parashikim: 1.71 cm në pajisjet mobile (pa kalibrim) 1.34 cm në pajisjet mobile (me kalibrim)	Të dhëna të mbledhura nga kamera e pajisjes mobile (paraqitja e imazheve)
TabletGaze (Huang, Veeraraghavan, & Sabharwal, 2015)	51 subjekte 4 qëndrime (angl. *postures)* 35 vendodhje vështrimi	Gabimi mesatar: 3.17cm	Të dhëna të mbledhura nga kamera e përparme e një pajisje tablet (paraqitja e imazheve)
TurkerGaze (Xu, et al., 2015)	3 pjesëmarrës 20.608 imazhe	Gabimi median: 1.06 gradë	Të dhëna të mbledhura nga kamera web (paraqitja e imazheve)
MPIIGaze (Zhang, Sugano, Fritz, & Bulling, 2015)	15 pjesëmarrës 213.659 imazhe	Gabimi mesatar i parashikimit: 13.9 gradë	Të dhëna të mbledhura nga kamera web (paraqitja e imazheve)
ETH-XGaze (Zhang, et al., 2020)	110 pjesëmarrës 1 milion kampione	Gabimi në gradë për detektimin e vështrimit: 4.2-5.4	Të dhëna të mbledhura nga 18 kamera digjitale SLR
WebQAmGaze (Ribeiro, Brandl, Søgaard, & Hollenstein, 2023)	332 pjesëmarrës	Saktësia: 70.3-73.4%	Të dhëna të mbledhura nga kamera ueb

Në Tabelën 3.7 janë listuar disa nga libraritë softuerike dhe algoritmet kryesore të cilat janë përdorur për të përmirësuar saktësinë e eksperimenteve të gjurmimit të syrit me pajisje me kosto të ulët. Papoutsaki et al. (2016) kanë prezantuar një librari Javascript të njohur me emërtimin WebGazer, e cila mundëson gjurmimin e vështrimit drejtpërdrejt në mjedise ueb. Kjo librari ofron një mekanizëm vetëkalibrimi të modelit, i cili bazohet në sjelljen e përdoruesve gjatë ndërveprimit me një faqe ueb. Të dhënat grumbullohen nëpërmjet përdorimit të kamerës ueb ose mobile. Kjo librari mund të integrohet lehtësisht në uebsajte me disa rreshta kodi dhe të mundësojë detektimin e bebes së syrit dhe të vlerësojë vendndodhjet e pikës së vështrimit duke përdorur analizën e regresonit (angl. *regression analysis*). Autorët kanë vlerësuar saktësinë e kësaj librarie nëpërmjet eksperimenteve online dhe mekanizmave të mësimit të vazhdueshëm nga ndërveprimi me përdoruesit.

Cheng, Ping, Wang dhe Chen (2022) kanë propozuar një metodë hibride për gjurmimin e syrit në pajisjet mobile ku autorët krahasojnë rezolucione të ndryshme të imazheve dhe të ndriçimit për të arritur rezultatet më të mira të mundshme. Ata prezantojnë metodën EasyGaze, e cila përdor imazhet e kapura nga një kamerë RGB (kamera standarde me kanalet e ngjyrave të kuqe, jeshile dhe blu) dhe kanë zhvilluar një studim me 12 pjesëmarrës duke përdorur pajisje smartphone me sistemin operativ Android. Metoda dallohet për disa avantazhe kryesore: performancë të qëndrueshme në kushte të ndryshme ndriçimi, funksionim të mirë edhe me imazhe me rezolucion shumë të ulët (deri në 26×13 piksel), si dhe implementim të thjeshtë në pajisje mobile.

Müller dhe Mann (2021) kanë prezantuar GazeClassify, një paketë softuerike me burim të hapur, e implementuar në Python për të etiketuar në mënyrë automatike të dhënat e mbledhura nga kamera e pajisjeve mobile. Algoritmi i tyre paraqet rezultate të mira krahasuar me identifikimin manual të vështrimit nga tre etiketues njerëzorë. Një tjetër platformë e rëndësishme është dhe Vocabulometer, një platformë ueb e prezantuar nga (Augereau, Jackuet, Kise, & Journet, 2018) e përshtatshme për t'u përdorur me Tobii 4C dhe Tobii Eye X, e cila mbledh të dhëna nga leximi i përdoruesve duke analizuar fjalët ku përqendrohet më shumë vështrimi dhe sjelljet e tyre gjatë leximit në kohë reale.

Park dhe Park (2016) kanë krijuar një algoritëm efikas për të detektuar beben e syrit në kohë reale në pajisjet për gjurmimin e syrit. Ndërkohë, metoda e propozuar nga (Hosp, et al., 2020), e njohur si një sistem me kosto të ulët për kapjen në kohë reale të vështrimit, synon vlerësimin e sistemeve me kosto të ulët për gjurmimin e lëvizjeve të syve. Autorët fokusohen në metodat për kapjen e lëvizjeve me shpejtësi të lartë. Ata kanë zhvilluar një softuer në C/C++**,** duke përdorur librari të njohura si OpenCV (për përpunimin e imazhit) dhe Qt (për ndërfaqen grafike) për të kapur imazhet, reflektimet në sy dhe detektimin e bebes së syrit. Sistemi i tyre ka një kosto totale prej 600 euro dhe arrin një nivel mesatar saktësie prej 0.38 gradësh. Modeli është i testuar në 19 pjesëmarrës dhe ka treguar saktësi prej 0.5 gradësh në rastin më të mirë dhe 1.4 gradë në rastin më të keq.

Tabela 3.7. Librari dhe sisteme për gjurmimin e syrit

Artikulli	**Emri**	**Karakteristika**
(Papoutsaki, et al., 2016)	WebGazer	Librari JavaScript
(Cheng, Ping, Wang, & Chen, 2022)	EasyGaze	Unity
(Muller & Mann, 2021)	GazeClassify	Paketë në Python
(Papoutsaki, Laskey, & Huang, 2017)	SearchGazer *(zgjerim i WebGazer)*	Librari JavaScript
(Augereau, Jackuet, Kise, & Journet, 2018)	Vocabulometer	Platformë Web
(Park & Park, 2016)	Sistem me kosto të ulët për kapjen në kohë reale të vështrimit	Pajisje Hardware; Driver; Software
(Hosp, et al., 2020)	Gjurmimi i lëvizjeve të syrit me shpejtësi të lartë	Pajisje Hardware C/C++, OpenCV
(Rakhmatulin & Duchowski, 2020)	Përdorimi i rrjeteve nervore të thelluara (angl. *Deep Neural Network)* për gjurmimin e syrit me kosto të ulët	Python, OpenCV
(Taban, Croock, & Korial, 2018)	Aplikacion mobile për gjurmimin e syrit	Android Studio, OpenCV

Rakhmatulin dhe Duchowski (2020) përdorin rrjetet nervore artificiale për të përmirësuar saktësinë e pajisjeve me kosto të ulët për gjurmimin e syrit. Autorët propozojnë një metodë për të kapur vështrimin, duke vendosur theksin në nevojën e pajisjeve me kosto të ulët dhe duke u përqendruar në nevojën e krijimit të standardeve në fushën e gjurmimit të syrit, në mënyrë që të mund të arrihet krahasueshmëria e rezultateve mes punimeve të ndryshme.

Së fundmi, Taban, Croock dhe Korial (2018) kanë krijuar një aplikacion mobile i cili mundëson përzgjedhjen e butonave në ndërfaqe në varësi të lëvizjes së syve. Ky aplikacion është ndërtuar duke përdorur Android Studio, mjedisin zyrtar për zhvillimin e aplikacioneve Android, dhe librarinë OpenCV e cila mundëson përpunimin e imazheve dhe analizën e të dhënave vizuale në kohë reale. Pas testimit të aplikacionit në një tablet Huawei, saktësia e arritur e modelit ishte 66% në mjediset e jashtme dhe 76% në mjediset e brendshme.

Përveç librarive dhe sistemeve të përmbledhura në Tabelën 3.7 autorë të shumtë e kanë përqendruar punën e tyre kërkimore në nënfusha të ndryshme si përzgjedhja e veçorive (angl. *feature selection*), detektimi i bebes së syrit (angl. *pupil detection*), kalibrimi (angl. *calibration*), parashikimi i gabimeve (angl. *error prediction*) dhe ndërtimi i pajisjeve me kosto të ulët, të paraqitura në Tabelën 3.8.

Tabela 3.8. Nënfusha kërkimi në gjurmimin e syrit me pajisje me kosto të ulët

Fusha e kërkimit	**Artikujt**
Përzgjedhja e veçorive (angl. *features*)	(Aunsri & Rattarom, 2022)
Detektimi i bebes së syrit	(Gomez-Poveda & Gaudioso, 2016)
Kalibrimi	(Lei, 2021), (Li, et al., 2022), (Khonglah & Khosla, 2015)
Parashikimi i gabimeve	(Wisiecka, et al., 2022), (Saxena, Lange, & Fink, 2022)
Ndërtimi i pajisjeve me kosto të ulët	(Wang, Zeng, & Liu, 2016), (Changwani & Sarode, 2019), (Mounica, Manvita, Jyotsna, & J., 2019), (Luo, et al., 2019), (Eide, et al., 2019), (Ivanchenko, Rifai, Hafed, & Schaeffel, 2020)

Në punën e tyre kërkimore Aunsri dhe Rattarom (2022) kanë propozuar përzgjedhjen e një bashkësie të re me veçori të gjurmimit të syrit në mënyrë që të arrihet një performancë më e mirë në aplikacionet në kohë reale dhe në sistemet me kosto të ulët. Veçoritë e përzgjedhura prej tyre përfshijnë lëvizjet e kokës dhe pozicionet e ndryshme të syve në lidhje me ekranin. Efikasiteti i veçorive të përzgjedhura, u testua nga autorët duke përdorur disa algoritme të mësuarit të makinës (angl. *machine learning*), duke arritur një nivel njohje prej 93.94%.

Ndërkohë Gómez-Poveda dhe Gaudioso (2016), kanë propozuar një njësi matëse për të vlerësuar stabilitetin e detektimit të bebes së syrit. Autorët kanë prezantuar këtë njësi duke përcaktuar se implementimi i saj në algoritmet ekzistuese do të rriste ndjeshëm saktësinë e detektimit. Kjo metrikë mund të përdoret për të krahasuar stabilitetin kohor të algoritmeve të ndryshme për gjurmimin e syrit.

Punime të tjera janë të ndërlidhura me procesin e kalibrimit, si një nga hapat më të rëndësishëm për të garantuar saktësinë e të dhënave të mbledhura. Autorët përshkruajnë si probleme kyç distancën mes përdoruesit dhe pajisjes, kushtet e pafavorshme të ndriçimit, rezolucionin e ulët të kamerës dhe faktorë të tjerë mjedisorë që ndikojnë drejtpërdrejt në saktësinë e vlerësimit të vështrimit.

Në këtë kontekst, Li et al. (2022) në punën e tyre kërkimore kanë studiuar gabimet e ndryshme të kalibrimit që ndodhin gjatë përcaktimit të vështrimit. Autorët i kanë studiuar këto gabime duke marrë në konsideratë veçoritë individuale në fytyrën e

pjesëmarrësve dhe duke përdorur një qasje probabilistike të bazuar në metodën Monte Carlo (metodë statistikore për simulimin e pasigurisë). Ndërkohë, Khonglah dhe Khosla (2015) kanë prezantuar një teknikë të bazuar në rreze infra të kuqe (angl. *infrared*) e cila mundëson ndërtimin e një sistemi me kosto të ulët i cili përdor një kamera të thjeshtë ueb, pa nevojën e kalibrimit. Rezultatet e këtij studimi synojnë aplikime në detektimin e hershëm të autizmit te fëmijët e vegjël.

Wisiecka et al. (2022) kanë zhvilluar një studim për të krahasuar gabimet e parashikimit në gjurmimin e syrit përmes kamerave ueb dhe sisteme profesionale të gjurmimit të syrit në distancë (angl. *remote eye-tracking*). Ata krahasojnë një zgjidhje të bazuar në kamerë ueb, përmes softuerit online RealEye, me pajisjen profesionale GP3, duke përdorur platformën eksperimentale PsychoPy. Në këtë eksperiment janë përfshirë 83 studentë dhe rezultatet finale treguan se gjurmimi i syrit duke përdorur kamerën ueb është një alternativë e pranueshme për aplikime kërkimore.

Saxena, Lange dhe Fink (2022) kanë propozuar përdorimin e metodave të të mësuarit të thelluar (angl. *deep learning)* për zhvillimin e eksperimenteve për gjurmimin e syrit online. Ata arritën në përfundimin se kombinimi i kalibrimit të pajisjes, kalibrimit të vështrimit dhe përfshirja e modelit të të mësuarit të thelluar sjell një përmirësim në saktësi prej 2.58 gradësh, në krahasim me eksperimentet pa përdorimin e modelit të të mësuarit të thelluar.

Punime të tjera prezantojnë forma të ndryshme të zhvillimit të eksperimenteve të gjurmimit të syrit me kosto të ulët. Shembuj të tillë përfshijnë punën e (Wang, Zeng, & Liu, 2016) ku autorët prezantojnë një sistem me kosto të ulët (38 USD), i cili vjen në formën e syzeve. Gabimi mesatar i gjeneruar nga sistemi i tyre është 0.5748 gradë. Changwani dhe Sarode (2019) kanë propozuar një sistem me kosto 50 USD, të përbërë nga dy kamera, të cilat kapin lëvizjen e syrit dhe të kokës. Me përfshirjen e rrjeteve nervore dhe vizionit kompjuterik (angl. *computer vision)* autorët kanë arritur të parashikojnë pikën ku po shikon pjesëmarrësi.

Mounica et al. (2019) kanë propozuar një sistem ku përdoren kamerat ueb dhe libraritë e vizionit kompjuterik, ndërsa Luo et al. (2019) zhvillojnë eksperimentin e tyre me 15 pjesëmarrës duke përdorur Tobii Eye Tracker 4C, me një kosto të përafërt prej 169 USD, duke arritur një saktësi prej 97.85%. Ata përdorin një klasifikues SVM (angl. *Support Vector Machine*) për të arritur një rezultat të tillë. Eide et al. (2019) krahasojnë një pajisje me kosto të ulët EyeX, me një pajisje me kosto më të lartë TX200 për të detektuar problemet *okulomotore* (lëvizje të syve) në 39 fëmijë duke arritur në rezultatin se dy pajisjet, pavarësisht ndryshimit në çmim arrijnë rezultate të ngjashme.

Së fundmi, Ivanchenko et al. (2020) prezantojnë një pajisje të ndërtuar prej tyre me karakteristika kryesore: kalibrimin e automatizuar, matjen e zhurmave, detektimin e pulitjeve dhe analizën e qendrës së bebes së syrit. Autorët e vlerësuan pajisjen e tyre me 10 pjesëmarrës, duke arritur në përfundimin që kjo pajisje mund të përdorej për gjurmimin e të dy syve gjatë eksperimenteve të ndryshme.

KREU 4

GJURMIMI I KURSORIT TË MAUSIT

4.1. Hyrje

Përdorimi i pajisjeve për gjurmimin e syrit luan një rol të rëndësishëm në kuptimin e sjelljeve gjatë leximit dhe proceseve gjuhësore si etiketimi i pjesëve të ligjëratës dhe lidhjeve sintaksore. Megjithatë, këto teknologji janë të ndërlidhura me kosto të larta financiare dhe kërkojnë pajisje të specializuara laboratorike, të cilat shpesh nuk janë të disponueshme për komunitetet kërkimore që merren me gjuhë të nënpërfaqësuara në burime digjitale, si gjuha shqipe. Këto kufizime financiare dhe infrastrukturore e vështirësojnë ndjeshëm kërkimin shkencor për kuptimin e sjelljeve gjatë të lexuarit për gjuhën shqipe dhe jo vetëm. Për të trajtuar këtë mangësi në këtë kapitull hulumtojmë përdorimin e alternativave me kosto të ulët për të studiuar sjelljen gjatë leximit në shqip. Në mënyrë specifike, fokusi vendoset në grumbullimin e të dhënave gjatë leximit përmes gjurmimit të lëvizjeve të kursorit të mausit (angl. *mouse tracking*), një metodë që mundëson mbledhjen e të dhënave në mënyrë të shkallëzueshme dhe pa nevojën e pajisjeve të specializuara.

Gjurmimi i kursorit të mausit është një alternativë praktike dhe ekonomike ndaj gjurmimit të syrit, veçanërisht në kontekste kërkimore ku qasja në pajisje profesionale është e kufizuar. Edhe pse lëvizjet e mausit nuk përfaqësojnë drejtpërdrejt vështrimin vizual, studime të ndryshme (Wilcox, Ding, Sachan, & Jäger, 2024) kanë treguar se ekziston një korrelacion domethënës midis trajektores së kursorit dhe shpërndarjes së vëmendjes konjitive gjatë leximit.

Eksperimentet e përshkruara në këtë pjesë kanë shërbyer si bazë për krijimin e bashkësisë së parë të të dhënave nga eksperimentet e gjurmimit të kursorit të mausit gjatë leximit në gjuhën shqipe. Kjo bashkësi e të dhënave mundëson analizën e sjelljes gjatë leximit të folësve amtarë të gjuhës shqipe, si dhe integrimin e mëtejshëm të metrikave të gjeneruara të leximit në modele të inteligjencës artificiale, duke përfaqësuar një kontribut për zhvillimin dhe përmirësimin e modeleve të gjuhësisë kompjuterike (angl. *computational linguistics)* për të parashikuar pjesët e ligjëratës (PoS), lidhjet sintaksore (angl. *syntactic dependencies)* etj.

Objektivat kryesore që kërkohen të arrihen nga zhvillimi i eksperimenteve për gjurmimin e kursorit të mausit janë:

1. Krijimi i një korpusi me të dhëna nga gjurmimi i kursorit të mausit gjatë leximit të teksteve në gjuhën shqipe.
2. Analizimi i të dhënave të mbledhura për të kuptuar sjelljet gjatë leximit në folësit amtarë të shqipes.
3. Përcaktimi dhe dokumentimi i hapave për parapërpunimin e të dhënave nga eksperimentet e gjurmimit të mausit dhe gjenerimi i metrikave të leximit.

4. Përcaktimi i një metodologjie për përdorimin e të dhënave nga gjurmimi i kursorit të mausit për të mbështetur përmirësimin e modeleve të PGJN-së për gjuhën shqipe me fokus nivelin e kuptimit gjuhësor.

Eksperimenti është i disponueshëm me burim të hapur në lidhjen https://github.com/Alba10/tracking dhe nuk kërkon pajisje të specializuara duke lehtësuar mbledhjen e të dhënave nga një grup më i gjerë lexuesish. Në këtë kapitull vendoset baza metodologjike mbi të cilën mund të ndërtohen analizat eksperimentale në kapitujt në vijim.

4.2. Veçoritë teknike të eksperimentit

Eksperimentet për gjurmimin e kursorit të mausit kanë si qëllim kryesor kapjen e pozicionit të kursorit të mausit dhe mbledhjen e kohës së leximit për secilën fjalë në tekst. Si rrjedhojë, për të zhvilluar eksperimentet e këtij tipi nuk kërkohen pajisje të avancuara, por vetëm një kompjuter i thjeshtë i lidhur me një pajisje maus. Ky hulumtim bazohet në studime të ngjashme të zhvilluara për gjuhën angleze përshkruar në nënkapitullin 2.4.

Në një eksperiment tipik për gjurmimin e mausit pjesëmarrësve u shfaqet një tekst, që fillimisht është jo i dukshëm (angl. *blurred)*, përveç një rajoni të vogël ku është e pozicionuar maja e kursorit të mausit. Dizajni i ndërfaqes së eksperimentit, i përbërë nga teksti jo mirë i dukshëm dhe rajone të fokusimit, është ideuar në mënyrë të tillë që të imitojë hapësirën vizuale të syrit të njeriut, me ndarjen mes *shikimit foveal* (angl. *foveal vision*, zona e shikimit me rezolucion të lartë) dhe *shikimit parafoveal* (angl. *parafoveal vision*, zona periferike me rezolucion më të ulët). Përdorimi i fshehjes së informacionit dhe i shfaqjes vetëm të një pjese të caktuar të tekstit për të studiuar vëmendjen vizuale në kushte eksperimentale është propozuar nga (Anwyl-Irvine, Armstrong, & Dalmaijer, 2022), por përveç studimit të vëmendjes, eksperimentet e gjurmimit të mausit mund të përdoren për të studiuar edhe përpunimin e gjuhës.

Para se të përfshihen në eksperiment, pjesëmarrësve u jepen udhëzime të qarta mbi mënyrën e ndërveprimit me ndërfaqen, duke theksuar se teksti shfaqet dhe lexohet vetëm përmes lëvizjes së kursorit të mausit. Në Figurën 4.1 paraqitet një shembull nga ndërfaqet e eksperimentit për gjurmimin e mausit, ku në fokus është fjala 'Hëna'. Ndryshe nga teknika e leximit me ritëm të vetëdrejtuar (angl. *self-paced reading*, SPR), gjurmimi i kursorit të mausit u mundëson pjesëmarrësve të kalojnë fjalë të caktuara pa i lexuar si edhe të kthehen prapa në pjesë të ndryshme të fjalisë. Ky fleksibilitet e bën procesin e leximit më afër leximit natyror, gjë e cila nuk është e mundur në teknikën SPR.

I gjithë eksperimenti mund të ekzekutohet në një shfletues interneti (angl. *browser)*, duke e kthyer atë në një eksperiment me kosto minimale për t'u zhvilluar. Pjesëmarrësit mund ta zhvillojnë eksperimentin nga shtëpia e tyre, pa nevojën e të qenit të pranishëm në një mjedis të kontrolluar laboratorik. Nuk kërkohen parametra të caktuar të kompjuterit, mjafton një ekran me përmasën 15 inç dhe përdorimi i një mausi. Rekomandohet që gjatë eksperimentit të mos përdoret paneli prekës i laptopit (angl.

touchpad). Një shembull i plotë ku mund të ekzekutohet i gjithë eksperimenti gjendet në lidhjen https://alba10.github.io/tracking/.

Pjesëmarrësit në eksperiment duhet të lëvizin mausin për të shfaqur dhe lexuar tekstin në rajone të ndryshme, duke zhvendosur pikën e fokusimit. Çdo lëvizje e mausit regjistrohet në mënyrë që më pas të përpunohet për të analizuar kohën e leximit për çdo fjalë dhe për të prodhuar *trajektoret e leximit* (angl. *scanpaths*), të cilat pasqyrojnë rrugën e ndjekur nga vëmendja gjatë leximit. I gjithë eksperimenti është implementuar nëpërmjet platformës Magpie (Ji, Klehr, Ilieva, Rautenstrauch, & Franke, 2025), një *kornizë softuerike* (angl. *framework*) e dedikuar për zhvillimin dhe ekzekutimin e eksperimenteve psikologjike dhe psikolinguistike në shfletues interneti.

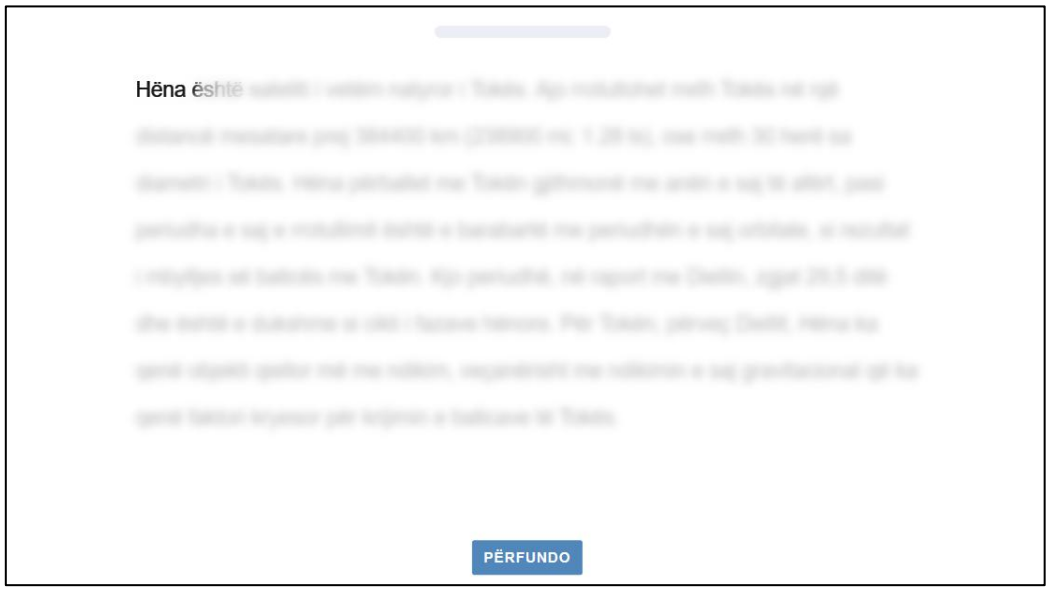


Figura 4.1. Ndërfaqja eksperimentale në eksperimentet e gjurmimit të kursorit të mausit

Teksti në eksperimentin e gjurmimit të mausit është paraqitur me shkrim Times New Roman, përmasa 12, hapësira dyshe (angl. *double-spaced)* midis rreshtave për t'u përafruar sa më shumë me eksperimentet e gjurmimit të syrit. Disa nga pikat kryesore në implementimin e eksperimentit janë: madhësia e zonës së fokusit (angl. *spotlight),* shikueshmëria e kursorit dhe niveli i dukshmërisë së fjalëve (angl. *bluriness)*. Madhësia e zonës së fokusit është një parametër kritik për interpretimin korrekt të sjelljes së leximit. Nëse zona e fokusit do të kishte përmasa të mëdha (p.sh. 4-5 fjalë), ose do të përfshinte disa rreshta të tekstit, atëherë do të ishte e vështirë të kuptohej se cilën fjalë po shikon pjesëmarrësi në eksperiment në një moment të caktuar. Në të kundërt, nëse zona e fokusit do të ishte shumë e vogël, pjesëmarrësit do të detyroheshin të bënin lëvizje shumë të vogla dhe të shpeshta gjë që mund të shkaktonte lodhje ose bezdi gjatë zhvillimit të eksperimentit. Për këtë arsye, duke u bazuar në argumentimin nga (Wilcox, Ding, Sachan, & Jäger, 2024), madhësia e zonës së fokusit është përzgjedhur në 102 pixel, duke synuar imitimin sa më të afërt të fokusit vizual të syrit të njeriut gjatë leximit.

Pikë tjetër kyçe është mënyra e paraqitjes dhe kapja e vendndodhjes së kursorit. Gjatë zhvillimit të eksperimentit kursori i mausit është i dukshëm në mënyrë që pjesëmarrësit të mos kenë vështirësi në lëvizjen e tij dhe në pozicionimin në tekst. Megjithatë, nëse kursori do të ishte i pozicionuar horizontalisht në qendër të zonës së fokusit ai do të shkaktonte vështirësi në leximin e fjalës. Për këtë arsye, kursori është i vendosur poshtë linjës horizontale të zonës së fokusit dhe pak i zhvendosur në të majtë të qendrës së zonës së fokusit, duke imituar faktin që zakonisht njerëzit gjatë leximit kapin pesë karakteret e para në të majtë të pikës së vështrimit dhe më pas pesëmbëdhjetë karakteret në të djathtë (McConkie & Rayner, 1975). Konkretisht, në këtë eksperiment, qendra funksionale e kursorit konsiderohet të jetë 39 pixel nga cepi në të majtë të zonës së fokusit dhe 63 pixel nga cepi i djathtë, duke u përshtatur me madhësinë e shkrimit dhe hapësirave midis rreshtave. Megjithatë, në varësi të tipit të studimit, këta parametra mund të ndryshohen përkatësisht në pjesën e implementimit në kod, për t'u përshtatur për tipa dhe madhësi të ndryshme shkrimi.

Një element shtesë në paraqitjen e zonës së fokusit është përdorimi i një tranzicioni gradual të dukshmërisë**,** ku qendra e fokusit është më e qartë, ndërsa zonat periferike paraqiten gjithnjë e më të paqarta. Ky efekt synon të imitojë rënien graduale të mprehtësisë vizuale në periferi të fushës së shikimit gjatë leximit natyror.

Në lidhje me frekuencën e kampionit dhe kapjen e pozicionit të mausit gjatë leximit është përdorur procedura e menaxhimit të të dhënave nga platforma Magpie. Gjatë eksperimentit, të dhënat dërgohen vazhdimisht nga ndërfaqja ueb në një server duke ruajtur për çdo regjistrim një *shënues kohor* (angl. *timestamp*) dhe numrin e identifikimit të pjesëmarrësit. Frekuenca e kampionimit e përzgjedhur për zhvillimin e eksperimentit është 20Hz, e cila korrespondon me regjistrimin e një mostre çdo 50 ms. Ky konfigurim gjeneron rreth 81 KB të dhëna për minutë duke mundësuar një ekzekutim të kënaqshëm të eksperimentit dhe duke shmangur problematika të lidhura me vonesat ose humbjen e të dhënave gjatë dërgimit të tyre në kohë reale.

4.3. Metodologjia e eksperimentit

Zhvillimi i eksperimentit të gjurmimit të mausit bazohet në mjetin MoTR (Wilcox, Ding, Sachan, & Jäger, 2024) dhe është përshtatur për zhvillimin e eksperimenteve tona në gjuhën shqipe në mënyrë që të jenë në koherencë me veçoritë specifike të studimit tonë. Metodologjia e eksperimentit përfshin disa hapa kryesorë të paraqitur në Figurën 4.2, dhe të shpjeguar në detaje më poshtë:

1. **Video shpjeguese:** Gjatë fazës së testimit, u vu re se pjesëmarrësit kishin vështirësi në kuptimin e procedurës dhe pritshmërive të eksperimentit, gjë e cila ndikonte në gatishmërinë e tyre për të marrë pjesë në eksperiment. Meqenëse eksperimenti nuk zhvillohet në kushte laboratorike, por shpërndahet në mënyrë elektronike dhe realizohet nga pjesëmarrësit në kompjuterët e tyre personalë, u krijua një video e shkurtër demonstruese. Qëllimi i saj është të sqarojë hapat kryesorë të eksperimentit, të reduktojë gabimet gjatë ekzekutimit dhe të rrisë përfshirjen e pjesëmarrësve.

2. **Ekrani i shpjegimeve:** Pas videos, pjesëmarrësve u shfaqet një ekran informues me shpjegime dhe disa udhëzime bazë që ata duhet të kenë parasysh gjatë zhvillimit të eksperimentit. Ky ekran përfshin informacion rreth mënyrës së ndërveprimit me tekstin, si dhe kohën mesatare të nevojshme për përfundimin e pjesës eksperimentale.
3. **Ekranet e leximit:** Pas ekranit të shpjegimeve, shfaqen njëri pas tjetrit të gjithë ekranet me tekstet përkatëse. Tekstet më të gjata janë të ndara në disa ekrane të njëpasnjëshme për të ruajtur lexueshmërinë dhe për të shmangur mbingarkesën vizuale.
4. **Pyetjet kuptimore:** Pjesëmarrësit duhet t'u përgjigjen pyetjeve të të kuptuarit pasi të kenë lexuar të gjitha ekranet që lidhen me një tekst të caktuar. Secila prej pyetjeve kuptimore ka dy alternativa të përgjigjes, ku vetëm njëra është e saktë. Në vijim kjo pjesë mund të zgjerohet me disa pyetje kuptimore, me më shumë se dy alternativa. Kjo do të mundësonte ruajtjen e më shumë informacioni në lidhje me kuptueshmërinë e tekstit, por për të ruajtur një kohë të arsyeshme të eksperimentit dhe për të mos rënduar pjesëmarrësit, numri i pyetjeve është mbajtur i kufizuar.
5. **Prezantimi i rezultateve:** Ekrani i fundit u prezanton pjesëmarrësve të dhënat e mbledhura gjatë kohës që ata zhvillojnë eksperimentin. Pjesëmarrësit kanë mundësinë t'i shkarkojnë të dhënat në kompjuterin e tyre në formatin *.csv.*
6. **Parapërpunimi i të dhënave:** Të dhënat e mbledhura përpunohen dhe analizohen për të përllogaritur metrikat e leximit, për të fshirë vlera të palejuara dhe për të krijuar një pamje vizuale për analizën e rezultateve të eksperimentit.

4.3.1. Përzgjedhja e teksteve dhe pyetjeve kuptimore

Eksperimenti kryesor përfshin tri tekste nga fusha të ndryshme, të gjitha të paraqitura në gjuhën shqipe. Tekstet janë zgjedhur në mënyrë të qëllimshme nga zhanre të ndryshme (letrar, argumentues dhe shkencor) me synimin për të mbledhur të dhëna sa më të larmishme mbi sjelljen e leximit. Duke zgjedhur tekste nga fusha të ndryshme, qëllimi ynë është të mbledhim të dhëna gjithëpërfshirëse dhe të analizojmë se çfarë ndikimi ka struktura dhe përmbajtja e një teksti në mënyrën se si një person lexon si edhe në nivelin e tij të përqendrimit.

Secili tekst është shoqëruar me një pyetje kuptimore, e cila shfaqet pas përfundimit të leximit dhe ka për qëllim të verifikojë nëse pjesëmarrësi ka ndjekur me vëmendje përmbajtjen e tekstit. Pyetjet kuptimore shërbejnë gjithashtu si mekanizëm kontrolli për cilësinë e të dhënave, duke ndihmuar në identifikimin e rasteve ku leximi mund të ketë qenë sipërfaqësor ose i pakujdesshëm.

Tri tekstet e përzgjedhura janë:

- *Hëna* - Një tekst enciklopedik i marrë nga Wikipedia, i cili përshkruan karakteristikat e satelitit natyror të Tokës. Ky tekst përmban disa paraqitje numerike dhe terminologji shkencore. Ai paraqitet në një ekran të vetëm, i ndjekur nga një pyetje kuptimore që i shfaqet pjesëmarrësit sapo përfundon së lexuari.

- *Prilli i Thyer* - Një fragment nga kapitulli i tretë i romanit të shkruar nga autori Ismail Kadare, në gjuhën origjinale shqipe. Teksti i përfshirë në eksperiment përmban një përshkrim të detajuar të peizazhit dhe të ndjesive të Besianit dhe Dianës, dy personazhet kryesore të romanit. Për shkak të gjatësisë së tij, ky tekst paraqitet me dhjetë ekrane të njëpasnjëshme dhe pas leximit të ekranit të dhjetë shfaqet një pyetje kuptimore.
- *Njeriu i Shpellave* - Një tekst shkencor i cili paraqet disa zbulime të rëndësishme arkeologjike nga një grup speleologësh. Ky tekst përmban terminologji shkencore dhe karakterizohet nga përdorimi i fjalive të gjata dhe komplekse. Teksti paraqitet në tetë ekrane të njëpasnjëshme, dhe pyetja kuptimore shfaqet pasi pjesëmarrësi përfundon së lexuari ekranin e tetë.

Një përmbledhje e karakteristikave për secilin prej teksteve të përfshira në eksperiment, duke përfshirë numrin e fjalëve dhe fjalive, si edhe pyetjen kuptimore përkatëse me alternativat dhe përgjigjen e saktë, paraqitet në Tabelën 4.1.

Tabela 4.1. Veçoritë e teksteve në eksperimentin e gjurmimit të kursorit të mausit

	HËNA	**PRILLI I THYER**	**NJERIU I SHPELLAVE**
Tipologjia e tekstit	Tekst enciklopedik	Roman	Tekst shkencor
Numri i fjalive	8	31	17
Numri i fjalëve	125	619	446
Pyetja kuptimore	Pse Hëna është kaq e rëndësishme për planetin Tokë?	Si u shpall në gazetë udhëtimi i Besian Vorpsit dhe gruas së tij?	Çfarë na tregon për skeletet datimi me karbon?
Alternativa 1	Sepse është i vetmi satelit i Tokës dhe ushtron forcë gravitacionale.	Në një njoftim privat që Besian Vorpsi i dërgoi një gazetari në Rrafsh.	Ato ishin nga koha përpara futjes së bujqësisë në Evropë.
Alternativa 2	Për shkak të masës së madhe dhe gravitetit sipërfaqësor të saj.	Si një muaj mjalti në male.	Ata i përkisnin burrave në të 30-at e tyre.
Përgjigjja e saktë	Sepse është i vetmi satelit i Tokës dhe ushtron forcë gravitacionale.	Si një muaj mjalti në male.	Ato ishin nga koha përpara futjes së bujqësisë në Evropë.

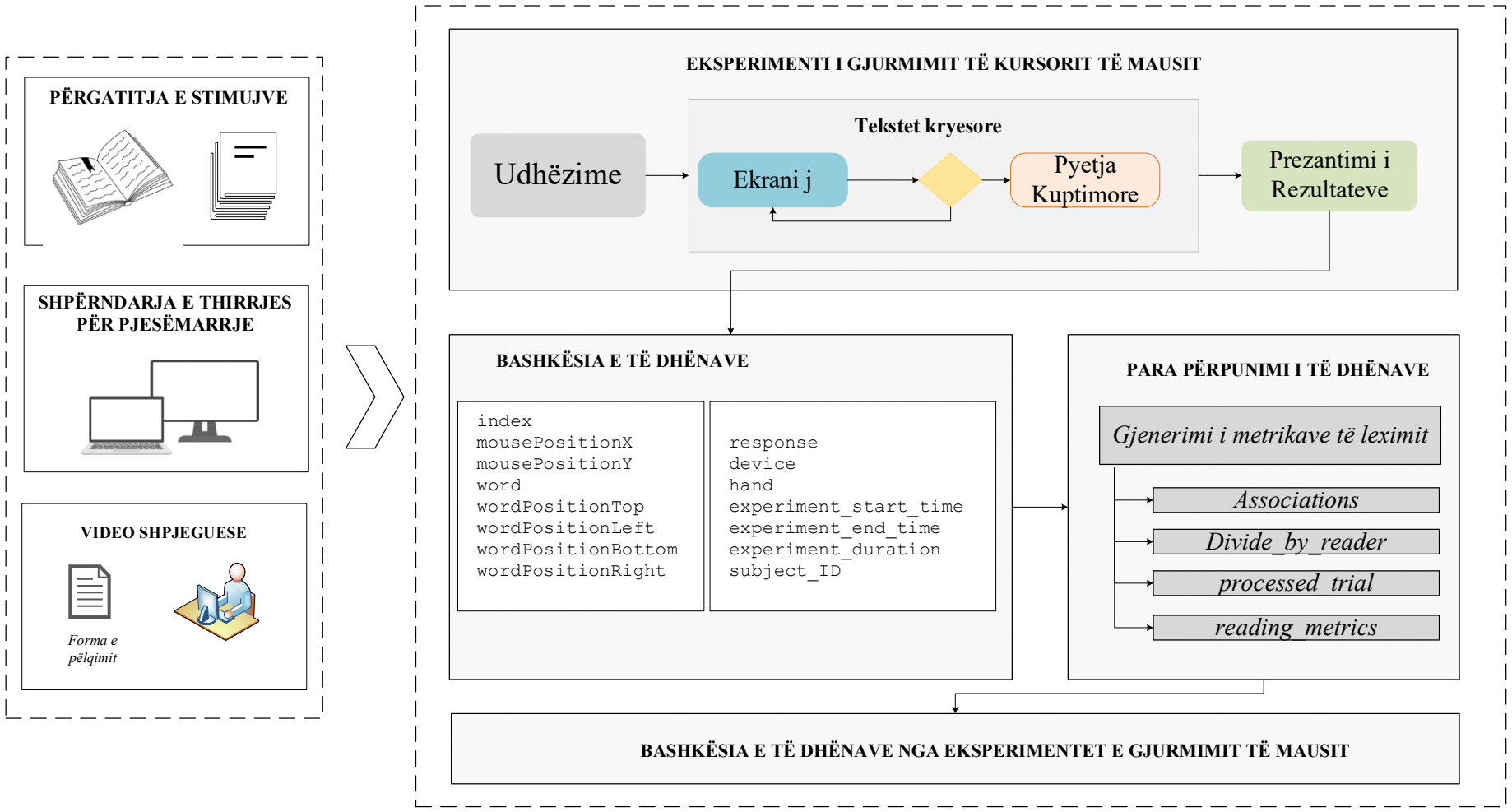


Figura 4.2. Arkitektura eksperimentale për studimin e gjurmimit të kursorit të mausit

Pyetjet kuptimore të paraqitura në fund të çdo teksti kanë si qëllim kryesor të matin nivelin e përqendrimit dhe kuptimit të tekstit nga secili pjesëmarrës në eksperiment. Në ndërfaqen e parë shpjeguese, e cila u paraqitet pjesëmarrësve para fillimit të eksperimentit, ata njoftohen se duhet të lexojnë tekstin me kujdes me qëllim që të kuptojnë përmbajtjen, gjë e cila do të kontrollohet me një pyetje kuptimore në fund të secilit tekst. Përfshirja e pyetjeve kuptimore shërben si një mekanizëm kontrolli për cilësinë e të dhënave, duke ndihmuar në ruajtjen e vëmendjes gjatë leximit dhe duke mundësuar filtrimin e mëvonshëm të rasteve ku pjesëmarrësit kanë dhënë përgjigje të pasakta. Kjo qasje siguron që të dhënat e gjurmimit të mausit të reflektojnë një proces real dhe të qëndrueshëm leximi, të lidhur drejtpërdrejt me përpunimin e përmbajtjes gjuhësore.

4.3.2. Bashkësia e të dhënave

Bashkësia e të dhënave të mbledhura nga eksperimenti i gjurmimit të kursorit të mausit përfshin 51 (pesëdhjetë e një) folës amtarë të gjuhës shqipe, 32 (tridhjetë e dy) studentë në programin *Inxhinieri Informatike* në grupmoshën 18-24, 11 (njëmbëdhjetë) pjesëmarrës i përkisnin grupmoshës 25-30, ndërsa 8 (tetë) pjesëmarrës ishin mbi 30 vjeç. Për sa i përket shpërndarjes gjinore, 29 pjesëmarrës ishin femra dhe 22 meshkuj. Të dhënat e mbledhura përfshijnë të gjithë skedarët në formatin `.csv` ku janë ruajtur të gjitha pozicionet e mausit gjatë procesit të leximit. Këta skedarë ruajnë informacion të detajuar në lidhje me të gjithë lëvizjet e mausit në tekst, duke ofruar të dhëna të pasura mbi mënyrën se si pjesëmarrës të ndryshëm ndërveprojnë me tekstin gjatë leximit. Për secilin nga 51 pjesëmarrësit është krijuar një skedar i veçantë i identifikuar përmes një numri unik identifikues.

Në përfundim të eksperimenteve të gjurmimit të mausit, bashkësia e të dhënave të krijuara përmban disa veçori (angl. *features)* të cilat ruajnë informacionin e grumbulluar gjatë leximit për secilin pjesëmarrës. Një përshkrim i detajuar i ndryshoreve kryesore të përfshira në bashkësinë e të dhënave paraqitet në Tabelën 4.2.

Tabela 4.2. Ndryshoret e gjeneruara nga eksperimenti i gjurmimit të kursorit të mausit

NDRYSHORJA	**PËRSHKRIMI**
`subjectId`	Numri identifikues unik i secilit pjesëmarrës në eksperiment.
`itemId`	Numri identifikues i ekranit të tekstit që i është shfaqur përdoruesit. Vlerat e `itemId` lëvizin nga 1 në 19, ku `itemId=1` lidhet me ekranin e parë dhe të vetëm për tekstin 'Hëna', vlerat e `itemId` nga 2 në 11 përkojnë me ekranet e ndërlidhura me tekstin 'Prilli i Thyer' dhe vlerat e `itemId` nga 12 në 19 përkojnë me ekranet e tekstit 'Njeriu i Shpellave'.

index	Pozicioni i fjalës në tekst, për secilin nga tekstet e përfshira. Indeksi me vlerë -1 përcakton hapësirat boshe dhe -100 përcakton përgjigjen e pyetjeve kuptimore.
mousePositionX	Ruan pozicionin në boshtin horizontal ku është pozicionuar kursori i mausit.
mousePositionY	Ruan pozicionin në boshtin vertikal ku është pozicionuar kursori i mausit.
word	Ruan fjalën ku është pozicionuar kursori i mausit.
wordPositionTop	Secila fjalë kornizohet nga boshtet lart, poshtë, majtas dhe djathtas. Kjo ndryshore ruan pozicionin e boshtit të sipërm të fjalës.
wordPositionLeft	Secila fjalë kornizohet nga boshtet lart, poshtë, majtas dhe djathtas. Kjo ndryshore ruan pozicionin e boshtit të majtë të fjalës.
wordPositionBottom	Secila fjalë kornizohet nga boshtet lart, poshtë, majtas dhe djathtas. Kjo ndryshore ruan pozicionin e boshtit të poshtëm të fjalës.
wordPositionRight	Secila fjalë kornizohet nga boshtet lart, poshtë, majtas dhe djathtas. Kjo ndryshore ruan pozicionin e boshtit të djathtë të fjalës.
response	Ndryshore e cila merr vlerën e përgjigjes së pjesëmarrësit për secilën pyetje kuptimore.
device	Ruan llojin e pajisjes së përdorur për të zhvilluar eksperimentin ku pjesëmarrësi zgjedh midis tre opsioneve: maus, panel prekës, tjetër.
hand	Përcakton cilën dorë përdor pjesëmarrësi për të lëvizur pajisjen, të majtën, të djathtën ose të dyja.
experiment_start_time	Koha e fillimit të eksperimentit
experiment_end_time	Koha e përfundimit të eksperimentit
experiment_duration	Kohëzgjatja totale e eksperimentit në ms (millisekonda).
responseTime	Fraksioni i kohës që ruhet sipas frekuencës së kampionimit (angl. *sampling rate*).

4.4. Parapërpunimi i të dhënave

Të dhënat e mbledhura nga eksperimentet e gjurmimit të kursorit të mausit mund të përdoren për një larmi qëllimesh, si ndryshueshmëria e kohëzgjatjes së leximit, studimi i kohës së leximit, identifikimi i fjalëve në tekst të cilat tërheqin më shumë vëmendjen, kuptimin midis lidhjes së kompleksitetit të një fjale dhe kohëzgjatjes së leximit etj. Përmes parapërpunimit të të dhënave, bëhet e mundur nxjerrja e metrikave që reflektojnë proceset konjitive të lidhura me leximin, duke e kthyer këtë bashkësi të dhënash në një burim të vlefshëm për studime në fushën e edukimit, psikolinguistikës dhe linguistikës kompjuterike.

Të gjitha skriptet në Python të përdorura për hapat e parapërpunimit janë ekzekutuar në një ambient cloud Google Colab Pro+, i cili ofron një infrastrukturë të përshtatshme për analizën e të dhënave në shkallë të gjerë. Ambienti i ekzekutimit bazohet në Python 3.12.12 dhe versioni Pro+ lejon shfrytëzimin e njësive të përpunimit të ofruar nga platforma, duke përfshirë GPU apo TPU, dhe memorie të rritur RAM në varësi të sesionit. Të gjithë të dhënat dhe skriptet janë ruajtur dhe menaxhuar në Google Drive, duke lehtësuar versionimin, aksesin dhe riprodhueshmërinë e analizave.

Të dhënat fillestare të regjistruara nga pajisja janë në një format *.csv* dhe përfaqësojnë një rrjedhë të vazhdueshme regjistrimesh të pozicionit të kursorit në ekran. Përpara se këto të dhëna të mund të analizohen në mënyrë kuptimplote, ato duhet t'i nënshtrohen një sërë hapash parapërpunimi, që synojnë pastrimin, strukturimin dhe transformimin e tyre në veçori analitike. Në eksperimentet klasike të gjurmimit të syrit, veçoritë që nxirren nga të dhënat fillestare janë fiksimet dhe sakadat. Në rastin tonë për eksperimentet e gjurmimit të mausit, duke u bazuar në arritjet e prezantuara në (Wilcox, Ding, Sachan, & Jäger, 2024), veçoria që nxirret nga përpunimi i të dhënave do të përcaktohet si pikë përqendrimi (angl. *attention association)* duke qenë se pjesëmarrësit gjatë zhvillimit të eksperimentit më së shumti prodhojnë lëvizje të ngadalta përgjatë tekstit. Përcaktimi i pikave të përqendrimit realizohet duke u bazuar në koordinatat x dhe y të kursorit në ekran, të cilat lidhen me pozicionin hapësinor të fjalëve në tekst. Përmes kësaj lidhjeje bëhet i mundur atribuimi i kohës së qëndrimit të kursorit fjalëve përkatëse, duke krijuar bazën për llogaritjen e metrikave të leximit në nivel fjale.

Në Figurën 4.3 paraqitet një shembull i formatit fillestar të ruajtjes së të dhënave nga eksperimenti për 1 pjesëmarrës, ku ilustrohet mënyra se si ruhen koordinatat horizontale dhe vertikale të kursorit të mausit gjatë procesit të leximit. Ky format përbën pikën e nisjes për të gjitha fazat e mëtejshme të parapërpunimit dhe analizës.

Secili pozicion i kursorit lidhet me fjalën që ndodhet në atë pozicion në ekran, ose, në rastet përkatëse, me hapësirat boshe ndërmjet fjalëve apo me zonat jashtë dritares së leximit. Të gjitha pikat e regjistruara që korrespondojnë me të njëjtën fjalë grupohen, duke përcaktuar kështu kohëzgjatjen totale gjatë së cilës ajo fjalë ka qenë në qendër të *zonës së fokusit* (angl. *spotlight*). Grumbullimi i këtyre pikave mundëson përcaktimin e metrikave të leximit në nivel fjale. Nga rezultatet e (Wilcox, Ding, Sachan, & Jäger, 2024), duke qenë se koha e përqendrimit e përllogaritur gjatë eksperimenteve të gjurmimit të mausit rezulton më e gjatë se koha tipike e kapur gjatë eksperimenteve të gjurmimit të syrit, është përcaktuar një kohë minimale dhe maksimale për të filtruar

pikat e përqendrimit që ishin ose shumë të shkurtra ose shumë të gjata, duke eleminuar problematikat e shkaktuara nga mbledhja e të dhënave në eksperimentet e zhvilluara në ndërfaqe ueb.

	Index	Word	mousePositionX	mousePositionY	wordPositionTop	\
0	104.0	kryesor	576.0	393.0	391.162506	
8	15.0	mesatare	549.0	165.0	151.162506	
9	15.0	mesatare	552.0	161.0	151.162506	
10	15.0	mesatare	552.0	161.0	151.162506	
11	15.0	mesatare	552.0	161.0	151.162506	
...	...	...	...	...	...	
8257	49.0	në	971.0	233.0	231.162506	
8258	49.0	në	971.0	233.0	231.162506	
8259	49.0	në	971.0	233.0	231.162506	
8260	50.0	Finlandë.	991.0	236.0	231.162506	
8261	50.0	Finlandë.	1014.0	246.0	231.162506	

	wordPositionLeft	wordPositionBottom	wordPositionRight
0	532.287537	411.162506	596.300041
8	505.262512	171.162506	585.300011
9	505.262512	171.162506	585.300011
10	505.262512	171.162506	585.300011
11	505.262512	171.162506	585.300011
...	...	...	...
8257	961.600037	251.162506	986.625036
8258	961.600037	251.162506	986.625036
8259	961.600037	251.162506	986.625036
8260	986.625000	251.162506	1060.675003
8261	986.625000	251.162506	1060.675003

Figura 4.3. Bashkësia fillestare e të dhënave nga gjurmimi i kursorit të mausit

Përkatësisht vlerat e përzgjedhura gjatë përpunimit të të dhënave ishin 120 ms deri në 4000 ms, vlera të cilat mund të ndryshohen në varësi të problemit studimor apo për të kryer krahasime të mëtejshme. Gjithashtu, nga të dhënat pastrohen të gjithë zhurmat që gjenerohen gjatë kohës që përdoruesi e lëviz mausin jashtë zonës ku paraqitet teksti që duhet lexuar.

Pas përpunimit të të dhënave, veçoritë e përllogaritura të cilat do të përdoren për pjesën studimore në KREU 6 janë paraqitur bashkë me shpjegimin përkatës në Tabelën 4.3. Për secilën metrikë u përllogaritën statistika përmbledhëse standarde, konkretisht: mesatarja, mediani, devijimi standard, gabimi standard i mesatares, vlerat minimale, maksimale dhe numri total i vëzhgimeve ndër pjesëmarrësit. Ky bashkim dhe këto fusha përmbledhëse mundësojnë një përfaqësim të sjelljes gjatë leximit në nivel fjale ndër pjesëmarrësit e përfshirë në studim. Në fazën përfundimtare, bashkësia e të dhënave u normalizua dhe standardizua, me qëllim eliminimin e ndikimit të shkallëve të ndryshme matëse dhe rritjen e stabilitetit numerik të modeleve të mësimit të makinës që do të aplikohen në analizat e mëtejshme.

4.5. Analiza e sjelljes gjatë leximit të pjesëmarrësve

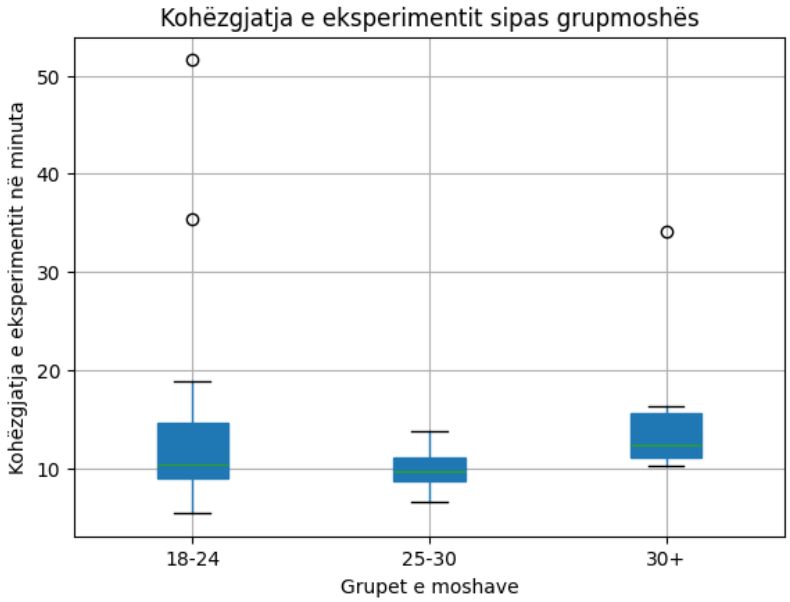
Analiza fillestare e të dhënave të mbledhura nga eksperimenti i gjurmimit të mausit përqendrohet në vlerësimin e kohëzgjatjes së eksperimentit për secilin nga pjesëmarrësit në mënyrë që të identifikohen vlera përjashtimore (angl. *outliers)*, si edhe për të matur kohën mesatare të përfundimit të eksperimentit sipas grupmoshave të ndryshme.

Tabela 4.3. Metrikat e përllogaritura nga leximi i teksteve në eksperimentet e gjurmimit të mausit

Veçori kryesore	
N fixations	Numri total i fiksimeve në një fjalë f.
Fixation probability	Probabiliteti për një fjalë f që të fiksohet.
Mean fixation duration	Mesatarja e kohës së të gjithë fiksimeve në një fjalë f.
Veçori fillestare	
first fixation duration	Kohëzgjatja e fiksimit të parë mbi një fjalë f.
Veçori të mëvonshme	
total fixation duration	Shuma e të gjithë fiksimeve në një fjalë f.
N re-fixations	Numri i herëve të rifiksimeve në një fjalë f (pas fiksimit të parë).
Re-read probability	Probabiliteti që një fjalë f të lexohet më shumë se një herë.
right_bounded_rt	Shuma e kohëzgjatjes së të gjithë fiksimeve në një fjalë f derisa fiksimi kalon në fjalën tjetër në të djathtë.
go_past_time	Shuma e të gjitha kohëzgjatjeve të përqendrimit duke nisur nga *first-pass association* në fjalë dhe deri sa përqendrimi kalon në një fjalë tjetër në të djathtë të fjalës.
Veçori të kontekstit	
w-2 fixation probability	Probabiliteti i fiksimit të fjalës para dy pozicioneve në fjali.
w-1 fixation probability	Probabiliteti i fiksimit të fjalës para një pozicioni në fjali.
w+2 fixation probability	Probabiliteti i fiksimit të fjalës pas dy pozicioneve në fjali.
w+1 fixation probability	Probabiliteti i fiksimit të fjalës pas një pozicioni në fjali.
w-2 fixation duration	Kohëzgjatja e fiksimit të një fjale para dy pozicioneve në fjali.
w-1 fixation duration	Kohëzgjatja e fiksimit të një fjale para një pozicioni në fjali.
w+2 fixation duration	Kohëzgjatja e fiksimit të një fjale pas dy pozicioneve në fjali.
w+1 fixation duration	Kohëzgjatja e fiksimit të një fjale pas një pozicioni në fjali.

Ky hap është i rëndësishëm për të siguruar cilësinë e të dhënave dhe për të kuptuar ndryshimet në kohën e përfundimit të eksperimentit sipas grupmoshave. Për të kuptuar shpërndarjen e kohëzgjatjes së eksperimentit në Grafikun 4.1 paraqiten 3 grupmosha: 18-24 (32 pjesëmarrës); 25-30 vjeç (11 pjesëmarrës); dhe mbi 30 vjeç (7 pjesëmarrës). Ajo që vërehet nga grafiku është fakti se pjesëmarrësit nga grupmosha të ndryshme kanë të njëjtën kohë mesatare të përfundimit të eksperimentit. Megjithatë, grupmosha 18–24 paraqet një ndryshueshmëri më të madhe në krahasim me dy grupet e tjera, si dhe dy vlera të dallueshme përjashtimore. Këto rezultate mund të sugjerojnë një ndryshim në nivelin e përqendrimit, fakt i cili lidhet me ndikimin më të madh që të rinjtë kanë ndaj shpërqendrimeve nga jashtë, duke sjellë pabarazi në kohëzgjatjen e eksperimentit.

Tre vlerat përjashtimore që vërehen në Grafikun 4.1 tregojnë se këta pjesëmarrës janë shpërqendruar gjatë eksperimentit dhe të dhënat e mbledhura gjatë gjurmimit të mausit nuk janë të sakta dhe nuk mund të përdoren në hapat e mëtejshëm të përpunimit. Për këtë arsye u vendos një vlerë limit (angl. *threshold)* si kufi i sipërm, ku secili pjesëmarrës që ka shpenzuar më shumë se 20 minuta për ta përfunduar eksperimentin nuk do të përfshihet në hapat e mëtejshëm të përpunimit. Kjo vlerë kufi u vendos duke u bazuar edhe në kohën mesatare të zhvillimit të eksperimentit, e cila për 50 pjesëmarrësit tanë është 12.7 minuta.

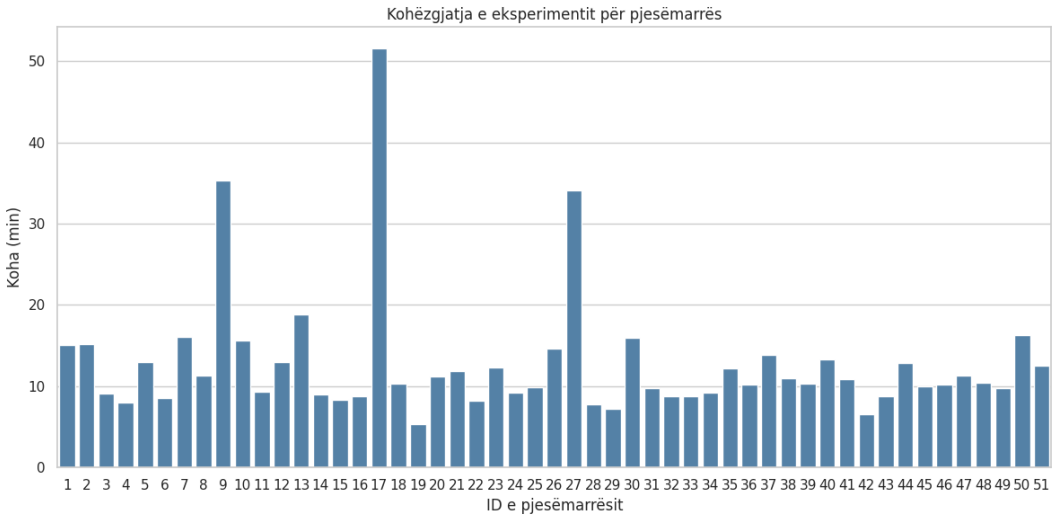


Grafiku 4.1. Kohëzgjatja e eksperimentit të gjurmimit të kursorit të mausit sipas moshës

Në Grafikun 4.2 paraqitet kohëzgjatja e eksperimentit për secilin nga pjesëmarrësit. Nga ky grafik shikojmë se kemi një shpërndarje relativisht të qëndrueshme ndërmjet pjesëmarrësve të ndryshëm (përveç vlerave përjashtimore). Mesatarja e kohëzgjatjes së eksperimentit është 12.77 minuta, me një devijim standard prej 7.76 minutash.

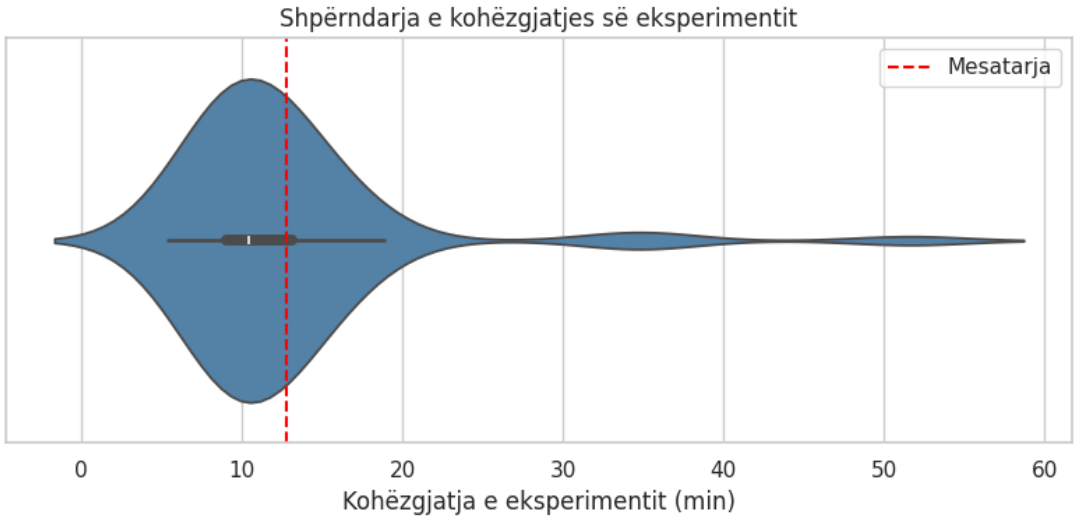
Pjesëmarrësi më i afërt me mesataren e kohëzgjatjes së eksperimentit është ai me numër identifikues 44, me një kohëzgjatje prej 12.88 minutash dhe 48 pjesëmarrës të

tjerë rezultojnë me vlera afër mesatares, që tregon një ritëm të ngjashëm leximi gjatë eksperimentit. Devijimin më të madh e shikojmë në pjesëmarrësin me id 17, ndjekur nga pjesëmarrësit me numër identifikues 9 dhe 27, të cilat nuk do të përfshihen në hapat e tjerë të përpunimit të të dhënave, pasi ka një probabilitet të madh që ato të mos kenë qenë të përqendruar gjatë eksperimentit.



Mesatarja	12.77 minuta	
Deviacioni standard	7.76 minuta	
Vlera më afër mesatares	*ID e pjesëmarrësit*	44
	Koha e leximit	12.88 minuta
	Ndryshesa nga mesatarja	0.11 minuta
Vlera më e largët nga mesatarja	*ID e pjesëmarrësit*	17
	Koha e leximit	51.61 minuta
	Ndryshesa nga mesatarja	38.84 minuta
Numri i pjesëmarrësve afër kohës mesatare	48	
Numri i pjesëmarrësve larg kohës mesatare	3	

Grafiku 4.2. Kohëzgjatja e eksperimentit të gjurmimit të mausit sipas pjesëmarrësve: mesatarja dhe ekstremet



Grafiku 4.3. Shpërndarja e kohëzgjatjes së eksperimentit të gjurmimit të mausit

Në Grafikun 4.3 vihet re qartë se shumica e kohëzgjatjeve janë të përqendruara rreth vlerës mesatare prej afërsisht dymbëdhjetë minutash, ndërsa disa vlera ekstreme krijojnë një bisht të djathtë në shpërndarje. Kjo formë shpërndarjeje përforcon nevojën për filtrimin e vlerave përjashtimore për të ruajtur qëndrueshmërinë dhe besueshmërinë e analizave të mëtejshme.

KREU 5

GJURMIMI I SYRIT PËRMES EYELINK PORTABLE DUO

5.1. Hyrje

Në këtë kapitull paraqitet në mënyrë të detajuar zhvillimi i eksperimenteve të gjurmimit të syrit gjatë leximit, të realizuara përmes pajisjes EyeLink Portable Duo. Kapitulli fokusohet në përshkrimin e plotë të konfigurimit teknik, metodologjisë eksperimentale, si dhe proceseve të mbledhjes dhe përpunimit fillestar të të dhënave.

Eksperimentet e përshkruara në këtë kapitull kanë shërbyer si bazë për krijimin e bashkësisë së parë të të dhënave të gjurmimit të syrit gjatë leximit për gjuhën shqipe. Kjo bashkësi të dhënash përbën një kontribut të rëndësishëm për zhvillimin e kërkimeve të mëtejshme për gjuhën shqipe në fushat e psikolinguistikës, shkencave konjitive dhe integrimit në modelet e përpunimit të gjuhës natyrore. Në mënyrë të veçantë, ajo mundëson analizën e sjelljes gjatë leximit në folës amtarë të gjuhës shqipe, si edhe integrimin e mëtejshëm të metrikave të gjeneruara të leximit në modele të inteligjencës artificiale për përpunimin e tekstit.

Një veçori e rëndësishme e këtij studimi është fakti se mbledhja e të dhënave është realizuar në përputhje me një iniciativë ku përfshihen shumë laboratorë në paralel, në kuadër të aksionit MultiplEYE (Jakobi, et al., 2025), i cili synon krijimin e një korpusi paralel shumëgjuhësh të gjurmimit të syrit gjatë leximit. Si rrjedhojë, gjatë gjithë procesit eksperimental është ndjekur një standard metodologjik i unifikuar i cili siguron krahasueshmërinë e të dhënave të mbledhura në gjuhën shqipe me të dhënat nga korpuset paralele për gjuhë të tjera.

Objektivat kryesorë që synohen të arrihen përmes zhvillimit të eksperimenteve për gjurmimin e syrit janë:

1. Krijimi i një bashkësie të standardizuar të të dhënave nga gjurmimi i syrit gjatë leximit të teksteve në gjuhën shqipe.
2. Analizimi i të dhënave të mbledhura për të kuptuar sjelljet gjatë leximit në folësit amtarë të shqipes.
3. Përcaktimi dhe dokumentimi i hapave për parapërpunimin e të dhënave dhe për gjenerimin e metrikave të leximit.
4. Hartimi i një metodologjie për përdorimin e të dhënave nga gjurmimi i syrit për të krijuar një bashkësi të dhënash e cila mund të mbështesë përmirësimin e modeleve të PGJN për gjuhën shqipe me fokus në informacionin gjuhësor dhe sjelljen gjatë leximit.

Në këtë kapitull vendoset baza metodologjike mbi të cilën ndërtohen analizat eksperimentale dhe integrimet në modelet e inteligjencës artificiale të trajtuar në kapitullin pasues.

5.2. Veçoritë teknike të eksperimentit

Pajisja e përdorur për zhvillimin e eksperimenteve të gjurmimit të syrit gjatë leximit është EyeLink Portable Duo (SR Research Ltd, 2016-2025), prodhuar nga SR Research EyeLink, një sistem profesional i përdorur gjerësisht në studime psikolinguistike dhe neurogjuhësore. Përshkrimi i detajuar i karakteristikave teknike të kësaj pajisjeje është dhënë më herët në nënkapitullin 2.2.2; në këtë seksion fokusohen vetëm parametrat konkretë të përdorur në eksperimentin tonë.

Zhvillimi i eksperimentit është realizuar duke ndjekur në mënyrë rigoroze protokollin e përcaktuar nga aksioni MultiplEYE (Jakobi, et al., 2025), me qëllim sigurimin e standardizimit metodologjik dhe krahasueshmërinë e të dhënave me korpuset paralele që janë duke u zhvilluar në laboratorë të ndryshëm ndërkombëtarë. Para nisjes së eksperimenteve janë përcaktuar disa kritere bazë teknike dhe procedurale, të cilat janë zbatuar në mënyrë të qëndrueshme gjatë gjithë procesit të mbledhjes së të dhënave.

Studimi është zhvilluar në mënyrë monokulare, duke regjistruar lëvizjet e syrit mbizotërues të secilit pjesëmarrës. Për këtë qëllim, në fillim të çdo sesioni eksperimental është kryer testi për përcaktimin e syrit mbizotërues në përputhje me standardet e rekomanduara në literaturë (MultiplEYE, 2024) dhe udhëzimet e protokollit MultiplEYE (Jakobi, et al., 2025). Frekuenca e kampionimit të pajisjes është përzgjedhur në 1000 Hz, që korrespondon me regjistrimin e 1000 kampioneve të pozicionit të syrit për sekondë. Kjo frekuencë lejon kapjen me një saktësi të lartë të lëvizjeve të shpejta të syrit, duke përfshirë fiksimet dhe sakadat, dhe është veçanërisht e përshtatshme për analizën e detajuar të sjelljeve të leximit në nivel fjale.

Për të minimizuar ndikimin e lëvizjeve të pavullnetshme të kokës dhe për të garantuar stabilitetin e regjistrimeve gjatë eksperimentit është përdorur një stabilizues koke. Përdorimi i stabilizuesit redukton prezencën e zhurmave në sinjalin e gjurmimit të syrit nga lëvizjet e kokës gjatë leximit. Distanca ndërmjet syrit mbizotërues të pjesëmarrësit dhe ekranit të kompjuterit është përcaktuar në 60 cm, në përputhje me rekomandimet standarde të pajisjes (SR Research Ltd, 2016-2025), ndërsa distanca ndërmjet syrit dhe pajisjes së gjurmimit të syrit është vendosur në 45 cm, duke siguruar një fushë optimale regjistrimi për kamerën infra të kuqe të sistemit të pajisjes EyeLink Portable Duo.

Gjatë procesit të kalibrimit dhe validimit të pajisjes, gabimi mesatar i pranuar i kalibrimit është mbajtur brenda kufirit maksimal prej 0.3 gradë këndore. Kalibrimi u realizua duke përdorur procesin e kalibrimit të pajisjes EyeLink Portable Duo me 9 pika (SR Research Ltd, 2016-2025). Vetëm në raste të rralla, kur ka qenë e pamundur të arrihet një gabim i tillë, kjo vlerë është tejkaluar, por gjithmonë duke mos shkuar mbi vlerën 0.5 gradë, në përputhje me kriteret e rekomanduara nga prodhuesit e pajisjes (SR Research Ltd, 2016-2025).

Të gjitha eksperimentet janë zhvilluar në një mjedis të kontrolluar laboratorik, duke përdorur një kompjuter me konfigurimet teknike të mëposhtme, i cili ka shërbyer si platformë për paraqitjen e stimujve dhe regjistrimin e të dhënave.

Specifikimet teknike të përdorura gjatë zhvillimit të eksperimenteve të gjurmimit të syrit me pajisjen EyeLink Portable Duo janë:

- Sistemi operativ: Windows 10
- Rezolucioni i monitorit (në piksel): 1920 x1080 px
- Madhësia fizike e ekranit (në centimetra): 54.3 x 30.2 cm
- Diagonalja e ekranit (në centimetra): 63 cm

Stimujt tekstual janë paraqitur duke përdorur një hapësirë (angl. *monospaced*) të shkronjave, dhe tip shkrimi JetBrainsMono-Regular dhe JetBrainsMono-ExtraBold. Zgjedhja e këtij tipi shkrimi siguron uniformitetin e hapësirës ndërmjet karaktereve dhe lehtëson përcaktimin e saktë të zonave të interesit *(Area of Interest – AOI)* në nivel karakteri dhe fjale. Për të shmangur lëvizje të papritura të tekstit gjatë eksperimentit dhe për të minimizuar gabime të mundshme në regjistrim, tekstet janë konvertuar në imazhe statike përpara paraqitjes së tyre në ekran.

5.3. Metodologjia e eksperimentit

Eksperimenti i gjurmimit të syrit gjatë leximit është organizuar në dy sesione kryesore. Sesioni i parë përfshin të gjithë hapat që lidhen me zhvillimin e eksperimentit të gjurmimit të syrit gjatë leximit, ndërsa sesioni i dytë përfshin zhvillimin e testeve psikometrike me qëllim vlerësimin e aftësive konjitive dhe të fjalorit të secilit pjesëmarrës. Mesatarisht sesioni i parë ka një kohëzgjatje prej dy deri në tre orë, duke përfshirë edhe pushimet e shkurtra të planifikuara, ndërkohë që sesioni i dytë ka një kohëzgjatje prej rreth një orë. Në varësi të disponueshmërisë së pjesëmarrësit, sesioni i dytë mund të zhvillohet në të njëjtën ditë ose në një ditë tjetër.

Eksperimenti është realizuar duke përdorur një algoritëm eksperimental të standardizuar (MultiplEYE-COST, 2025) i cili ekzekutohet në kompjuterin e paraqitjes së stimujve *(Display PC)*, ndërsa monitorimi i regjistrimeve është kryer nga kompjuteri i lidhur me pajisjen *(Host PC)*. Hapat kryesorë të metodologjisë së eksperimentit të paraqitur në Figurën 5.1 shpjegohen në detaje më poshtë:

1. **Përgatitja e mjedisit eksperimental:** Para nisjes së çdo sesioni verifikohet funksionimi korrekt i eksperimentit në kompjuterin e paraqitjes së stimujve (*Display_PC*) dhe kontrollohen parametrat e pajisjes së gjurmimit në kompjuterin pritës të pajisjes së gjurmimit të syrit (*Host_PC*). Për secilin pjesëmarrës caktohet një numër unik identifikues.
2. **Formulari i pëlqimit:** Pjesëmarrësi mirëpritet në laborator, informohet në lidhje me eksperimentin, nënshkruan formularin e pëlqimit ku i paraqiten qëllimi i studimit, procedura eksperimentale dhe të drejtat e pjesëmarrësit mbi të dhënat.
3. **Testi i syrit mbizotërues:** Kryhet testi për përcaktimin e syrit mbizotërues në mënyrë që pajisja të konfigurohet për regjistrimin monokular të syrit mbizotërues.
4. **Distancat eksperimentale:** Kontrolli i distancave të duhura midis pajisjes, ekranit dhe pjesëmarrësit sipas kushteve të paracaktuara.

5. **Ekranet e udhëzimeve:** Pjesëmarrësve u shfaqen ekranet fillestare të udhëzimeve, të cilat ofrojnë informacion bazë mbi mënyrën e zhvillimit të eksperimentit, si edhe mbi kohën mesatare të përfundimit (angl. *average completion time*) të pjesës eksperimentale.
6. **Kalibrimi dhe validimi:** Para fillimit të leximit kryhet procesi i kalibrimit dhe validimit të pajisjes, me qëllim sigurimin e saktësisë së gjurmimit të syrit.
7. **Ekranet e leximit:** Pas ekraneve të udhëzimeve shfaqen njëri pas tjetrit të gjithë ekranet me tekstet përkatëse. Tekstet më të gjata janë të ndara në disa ekrane. Duke qenë se eksperimenti ka një kohë relativisht të gjatë, pjesëmarrësit mund të lëvizin në pjesën e pushimeve të shkurtra midis teksteve, si edhe kanë një pushim të detyruar në mesin e eksperimentit. Pas çdo pushimi është e detyrueshme të rikryhet procesi i kalibrimit dhe validimit të pajisjes. Gjatë kohës që pjesëmarrësi lexon tekstet, eksperimentuesi mundet që të vështrojë lëvizjen e syrit të tij në *Host_PC*. Nëse vihen re zhvendosje të mëdha graduale të sinjalit *(*angl. *signal drift)* eksperimentuesi mund të kryejë sërish procesin e kalibrimit të pajisjes.
8. **Pyetjet kuptimore:** Pjesëmarrësit duhet ti përgjigjen dy pyetjeve kuptimore pas teksteve të praktikës dhe gjashtë pyetjeve kuptimore pas teksteve kryesore. Pyetjet kuptimore shoqërohen me katër alternativa të mundshme, nga të cilat vetëm njëra është e saktë.
9. **Pyetësori i pjesëmarrësve:** Në përfundimin e eksperimentit, pjesëmarrësit duhet të plotësojnë një pyetësor me disa pyetjeve demografike. Konkretisht pjesëmarrësit pyeten për gjininë, vitet e edukimit, nivelin e edukimit, moshën, statusin socio-ekonomik, gjuhën amtare, gjuhën e përdorimit, kohën që shpenzon mesatarisht duke lexuar në secilën gjuhë, nivelin e lodhjes, nëse përdorin syze apo lente në kohën e zhvillimit të eksperimentit si dhe nëse kanë konsumuar alkool atë ditë ose një ditë më parë.
10. **Testet psikometrike:** Sesioni i dytë përfshin administrimin e gjashtë testeve psikometrike. Ky sesion mund të zhvillohet në të njëjtën ditë me eksperimentin (të paktën tridhjetë minuta pas sesionit të parë të gjurmimit të syrit) ose në një ditë tjetër.
    - Kapaciteti i kujtesës **(**LWMC - Lewandosky Working Memory Capacity Battery) (Lewandosky, Oberauer, Yang, & K H Ecker, 2010)
    - Aftësia metagjuhësore (PLAB – Pimsleur Language Aptitude Battery) (Pimsleur, 1966)
    - Kontrolli konjitiv (Flanker - Cognitive control) (Eriksen & Eriksen, 1974)
    - Emërtimi i shpejtë i automatizuar (RAN - Rapid Automatized naming) (Denckla & Cutting, 1999)
    - Testi i fjalorit (WikiVocab/LexTale - Vocabulary test) (Lemhofer & Broersma, 2011)
11. **Parapërpunimi i të dhënave:** Të dhënat e mbledhura përpunohen dhe analizohen për të përllogaritur metrikat e leximit *(*angl. *reading metrics)*, për të filtruar vlera të palejuara dhe për të interpretuar rezultatet e eksperimentit.

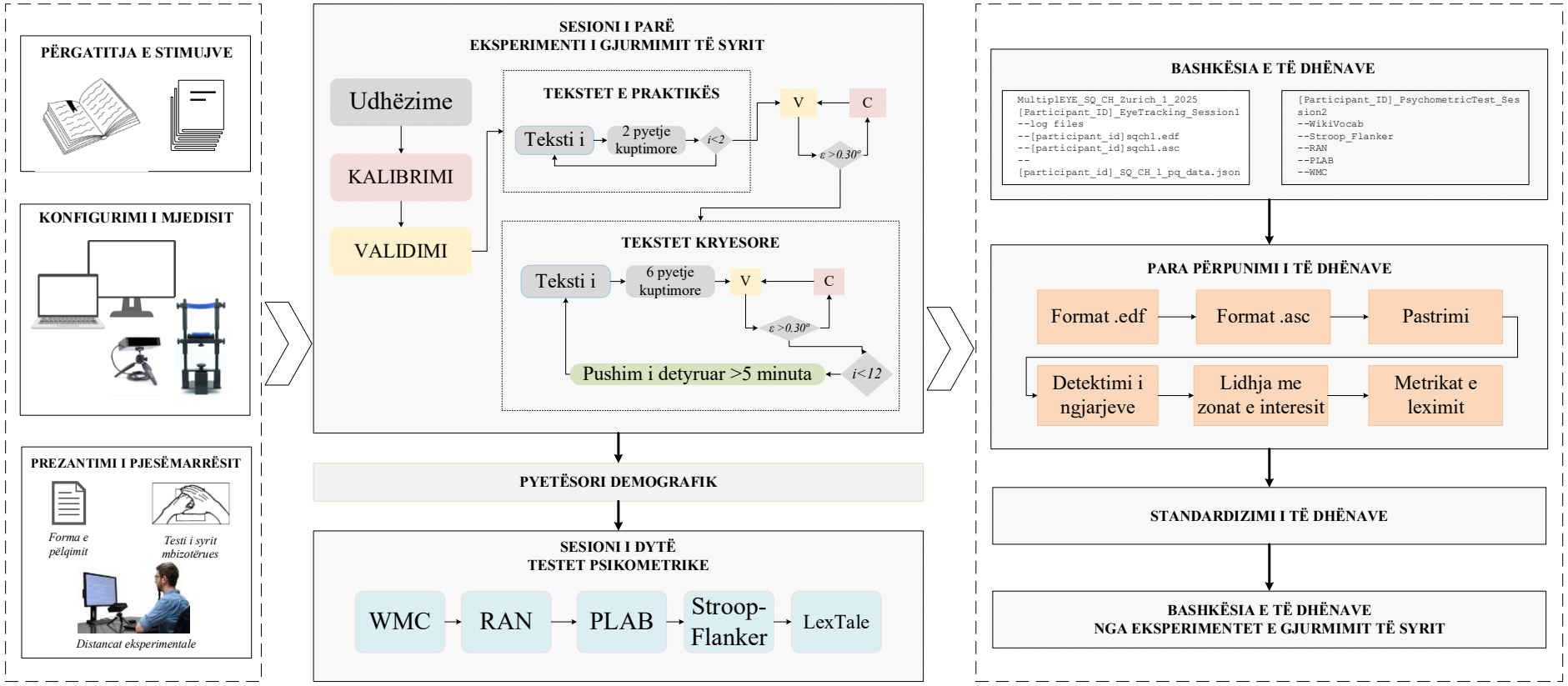


Figura 5.1. Arkitektura eksperimentale për studimin e gjurmimit të syrit

5.3.1. Përzgjedhja e teksteve dhe pyetjeve kuptimore

Eksperimenti kryesor përfshin gjithsej dymbëdhjetë tekste të përzgjedhura nga fusha të ndryshme, të gjitha të paraqitura në gjuhën shqipe. Këto tekste përfaqësojnë tipologji të ndryshme diskursive, duke përfshirë tekste letrare, shkencore, argumentuese dhe institucionale. Duke përzgjedhur tekste nga fusha të ndryshme, qëllimi ynë është të mbledhim të dhëna gjithëpërfshirëse dhe të analizojmë se çfarë ndikimi ka struktura dhe përmbajtja e një teksti në mënyrën se si një person lexon, si edhe në nivelin e tij të përqendrimit. Gjithashtu, kjo mund të ndihmojë në analizën e ndikimit që struktura sintaksore, kompleksiteti leksikor apo përmbajtja semantike e tekstit kanë në sjelljen gjatë leximit.

Tekstet janë ndarë në tekste praktike dhe tekstet kryesore eksperimentale, në përputhje me rolin e tyre në strukturën e eksperimentit.

Tekstet e praktikës

Tekstet e praktikës janë relativisht më të shkurtra dhe me një kompleksitet më të ulët gjuhësor. Ato shfaqen në fazën fillestare të eksperimentit dhe kanë për qëllim familjarizimin e pjesëmarrësit me procedurën eksperimentale, konfigurimin e ndërfaqes dhe mënyrën e leximit në kushtet e gjurmimit të syrit.

- *Hëna (The Moon*)
  Një tekst enciklopedik i marrë nga Wikipedia, i cili përshkruan karakteristikat e satelitit natyror të Tokës. Ky tekst përmban disa paraqitje numerike dhe terminologji shkencore. Ai paraqitet në një ekran të vetëm, i ndjekur nga dy pyetje kuptimore që i shfaqen pjesëmarrësit sapo përfundon së lexuari.

- *Era e Veriut* (*The North Wind*)
  Një tekst letrar i marrë nga përrallat e Ezopit, me origjinë nga Greqia e lashtë. Në eksperiment përfshihet përralla e plotë në një ekran të vetëm, e ndjekur nga dy pyetje kuptimore.

Tekstet kryesore

Tekstet kryesore përbëjnë pjesën qendrore të eksperimentit dhe karakterizohen nga një nivel më i lartë kompleksiteti gjuhësor dhe strukturor, duke përfshirë fjali më të gjata, terminologji të specializuar dhe struktura diskursive më të ndërlikuara.

- *Prilli i Thyer (Broken April*)
  Një fragment nga kapitulli i tretë i romanit të shkruar nga autori Ismail Kadare, me gjuhë origjinale gjuhën shqipe. Teksti i përfshirë në eksperiment përmban një përshkrim të detajuar të peizazhit dhe të ndjesive të Besianit dhe Dianës, dy personazhet kryesorë të romanit. Për shkak të gjatësisë së tij ky tekst paraqitet me dhjetë ekrane të njëpasnjëshme.

- *Njeriu i Shpellave* (*The Caveman*)
  Një tekst shkencor, i cili paraqet disa zbulime të rëndësishme arkeologjike nga një grup speleologësh. Ky tekst karakterizohet nga terminologji shkencore dhe nga përdorimi i fjalive të gjata e komplekse. Teksti paraqitet në tetë ekrane të njëpasnjëshme.

- *MultiplEYE*
  Një tekst shkencor i cili paraqet informacionin kryesor mbi projektin MultiplEYE. Ky tekst përfshin terminologji shkencore dhe ndahen në njëmbëdhjetë ekrane të njëpasnjëshme.

- *Deklarata e përgjithshme për të drejtat e njeriut* (*Universal Declaration of Human Rights*)
  Në eksperiment përfshihet hyrja e këtij dokumenti institucional të miratuar nga Organizata e Kombeve të Bashkuara. Ky tekst ndahet në tetë ekrane të njëpasnjëshme.

- *Mobiliteti për studime* (*Learning Mobility*)
  Në eksperiment përfshihet pjesa e parë e hyrjes së dokumentit e marrë nga raporti i progresit nga EUR-Lex, databaza e Ligjeve të Bashkimit Evropian. Teksti është i ndarë në tetë faqe të njëpasnjëshme.

- *Alkimisti* (*The Alchemist*)
  Një tekst letrar i marrë nga libri i Paulo Coelho, ku në eksperiment përdoren paragrafët e parë në kapitullin e parë. Teksti është i ndarë në pesë ekrane të njëpasnjëshme.

- *Mali Magjik* (*The Magic Mountain*)
  Një tekst letrar i marrë nga fjala hyrëse e librit 'Mali Magjik' i Tomas Man. Teksti është i ndarë në gjashtë ekrane të njëpasnjëshme.

- *Solaris* (*Solaris*)
  Një tekst letrar i shkruar nga Stanislav Lem. Për pjesën e eksperimentit është marrë një pjesë nga kapitulli i dytë, i ndarë në gjashtë ekrane të njëpasnjëshme.

- *Qumështi i lopës* (Cow's Milk)
  Një tekst argumentues i marrë nga testimet e PISA *(Programme for International Student Assessment)*. Për eksperimentin tonë është përzgjedhur një pjesë nga hyrja, e ndarë në katërmbëdhjetë faqe të njëpasnjëshme.

- *Rapa Nui (Rapa Nui)*
  Një tekst argumentues i marrë nga testimet e PISA *(Programme for International Student Assessment)*, i ndarë në trembëdhjetë ekrane të njëpasnjëshme.

Një përmbledhje e karakteristikave për secilin tekst, duke përfshirë numrin e fjalëve, numrin e fjalive dhe gjatësinë mesatare të fjalës dhe fjalisë, paraqitet në Tabelën 5.1.

Tabela 5.1. Veçoritë e teksteve në eksperimentin e gjurmimit të syrit

Teksti	**Tipologjia e tekstit**	**Numri i fjalive**	**Numri i fjalëve**	**Gjatësia mesatare e fjalisë**	**Gjatësia mesatare e fjalës**
Hëna	Tekst enciklopedik	8	125	15.62	4.49
Era e Veriut	Tekst enciklopedik	5	112	22.4	4.04
Prilli i Thyer	Literaturë	31	619	19.97	4.33
Njeriu i Shpellave	Tekst shkencor	17	446	26.24	5
MultiplEYE	Tekst shkencor	38	924	24.32	5.15
Deklarata e të drejtave të njeriut	Tekst institucional	1	366	366	4.93
Mobiliteti i mësimit	Tekst institucional	21	616	29.33	5.13
Alkimisti	Literaturë	25	468	18.72	4.2
Mali Magjik	Literaturë	13	529	40.69	4.2
Solaris	Literaturë	34	974	28.65	5.13
Qumështi i lopës	Tekst argumentues	58	923	15.91	5.04
Rapa Nui	Tekst argumentues	44	732	16.64	4.59

Pyetjet kuptimore të paraqitura në fund të çdo teksti kanë si qëllim kryesor vlerësimin e nivelit të përqendrimit dhe të kuptimit të përmbajtjes tekstuale nga secili pjesëmarrës gjatë eksperimentit. Përfshirja e pyetjeve kuptimore nxit përqendrimin e

lexuesit dhe ofron një mundësi për identifikimin dhe filtrimin e pjesëmarrësve të cilët demonstrojnë një nivel minimal të kuptimit të tekstit.

5.3.2. Bashkësia e të dhënave

Bashkësia e të dhënave e krijuar nga eksperimenti i gjurmimit të syrit përfshin 25 pjesëmarrës folës amtarë të gjuhës shqipe. Eksperimentet janë zhvilluar në Laboratorin e Linguistikës Digjitale të Universitetit të Zyrihut, me pjesëmarrës me origjinë shqiptare të cilët ishin rritur ose zhvendosur në Zyrih. Duke marrë parasysh vështirësinë në identifikimin e një numri të madh folësish amtarë të shqipes në këtë qytet, u vendos që në studim të përfshihen edhe pjesëmarrës që e kanë përvetësuar gjuhën shqipe në një mjedis familjar, dhe jo në mënyrë të strukturuar si pjesë e sistemit arsimor.

Të dhënat e mbledhura janë organizuar në mënyrë të strukturuar, ku për secilin pjesëmarrës është krijuar një dosje individuale e identifikuar me kodin unik të pjesëmarrësit, në përputhje me parimet etike të mbrojtjes së të dhënave personale. Çdo dosje përmban skedarët e krijuar gjatë eksperimentit të gjurmimit të syrit, të dhënat e stimujve dhe informacionin e mbledhur nga pyetësori demografik. Dosja në përfundim të sesionit të parë dhe të dytë për një pjesëmarrës ka strukturën e mëposhtme:

```
MultiplEYE_SQ_CH_Zurich_1_2025
[Participant_ID]_EyeTracking_Session1
--log files
-----Experiment_LOGFILE.txt
-----General_LOGFILE.txt
-----Stimuli_order.csv
-----Question_order.csv
--[participant_id]sqch1.edf
--[participant_id]sqch1.asc
--[participant_id]_SQ_CH_1_pq_data.json

[Participant_ID]_PsychometricTest_Session2
--WikiVocab
-----SQCH1_[participant_id]_PT2_timestamp.csv
--Stroop_Flanker
-----SQCH1_[participant_id]_PT2_timestamp.csv
-----SQCH1_[participant_id]_PT2_timestamp.psydat
--RAN
-----audio_SQCH1_[participant_id]_PT2
--------SQCH1_[participant_id]_S2_trial1.wav
--------SQCH1_[participant_id]_S2_trial2.wav
-----SQCH1_[participant_id]_PT2_timestamp.csv
--PLAB
-----SQCH1_[participant_id]_timestamp.csv
-----SQCH1_[participant_id]_timestamp.psydat
```

```
--WMC
-----SQCH1_[participant_id]_PT1_timestamp.csv
-----SQCH1_[participant_id]_PT1_timestamp.psydat
```

5.4. Parapërpunimi i të dhënave

Për parapërpunimin e të dhënave për secilin nga pjesëmarrësit dhe për çdo tekst janë ndjekur disa hapa të strukturuar për të arritur nga formati binar i regjistruar nga pajisja EyeLink Portable Duo deri në gjenerimin e metrikave kryesore të leximit, të cilat mund të përdoren për analiza të mëtejshme. Në këtë proces janë përfshirë përdorimi i mjeteve dhe algoritmeve specifik për trajtimin e të dhënave të gjurmimit të syrit.

Të gjitha skriptet në Python për hapat e parapërpunimit janë ekzekutuar në një ambient cloud në platformën Google Colab Pro+, i cili ofron një mjedis të përshtatshëm për përpunimin dhe analizën e të dhënave. Ambienti i ekzekutimit është bazuar në Python 3.12.12. Versioni Pro+ lejon shfrytëzimin e njësive të përpunimit të ofruar nga platforma, duke përfshirë GPU apo TPU, dhe memorie të rritur RAM në varësi të sesionit. Të gjithë të dhënat dhe skriptet janë ruajtur dhe menaxhuar në Google Drive, duke siguruar një akses më të lehtë gjatë eksperimentimit.

Të dhënat fillestare të regjistruara nga pajisja janë në një format binar të palexueshëm *.edf*. Për t'i bërë të dhënat të lexueshme ato janë konvertuar në formatin *.asc* duke përdorur mjetin edf2asc, i cili është pjesë e softuerit të EyeLink nga SR Research. File-i i gjeneruar *.asc* përmban të dhënat për një sesion të plotë, duke përfshirë metadata, dhe ka shërbyer si bazë për nxjerrjen e të dhënave specifike për secilin tekst të përfshirë në eksperiment. Ky file është përpunuar më tej për të kapur të dhënat për secilin tekst që përfshihet në eksperiment. Në këtë file, çdo rresht përmban një kampion me koordinatat x dhe y të fokusit të syrit dhe shënuesin kohor *(timestamp)* kur është regjistruar ky kampion.

Për secilin pjesëmarrës janë gjeneruar disa elemente kryesore:

- Një skedar binar *.edf* që përmban të gjithë lëvizjet e syrit dhe mesazhet përcaktuese për ngjarjet e ndryshme të ndodhura gjatë eksperimentit.
- Një skedar *.json* ku është ruajtur informacioni i marrë nga pyetësori.
- Disa *log_file* të cilat paraqesin të dhënat bazë mbi renditjen rastësore të teksteve dhe pyetjeve kuptimore.

Në Figurën 5.2 paraqitet një shembull i formatit ASCII duke treguar disa nga fushat kryesore të cilat mblidhen gjatë eksperimentit për secilin nga pjesëmarrësit dhe mënyrën se si janë të ruajtura të gjitha pozicionet e lëvizjes së syrit.

```
MSG 2031819 start_recording_PRACTICE_trial_1_stimulus_Enc_WikiMoon_13_page_1
MSG 2031824 RECCFG CR 1000 0 0 L
MSG 2031824 ELCLCFG BTABLER
MSG 2031824 GAZE_COORDS 0.00 0.00 1308.00 1001.00
MSG 2031824 THRESHOLDS L 67 228
MSG 2031824 ELCL_WINDOW_SIZES 176 188 0 0
MSG 2031824 CAMERA_LENS_FOCAL_LENGTH 27.00
MSG 2031824 PUPIL_DATA_TYPE RAW_AUTOSLIP
MSG 2031824 ELCL_PROC CENTROID (3)
MSG 2031824 ELCL_PCR_PARAM 5 3.0
MSG 2031825 !MODE RECORD CR 1000 0 0 L

START   2031827     LEFT    SAMPLES EVENTS
PRESCALER   1
VPRESCALER  1
PUPIL   AREA
EVENTS  GAZE    LEFT    RATE    1000.00 TRACKING    CR  FILTER  0
SAMPLES GAZE    LEFT    RATE    1000.00 TRACKING    CR  FILTER  0   INPUT
INPUT   2031827 0
2031827    75.2    92.2   836.0     0.0 ...
2031828    76.1    93.6   830.0     0.0 ...
2031829    73.9    92.6   829.0     0.0 ...
2031830    74.5    92.2   832.0     0.0 ...
2031831    75.0    92.6   824.0     0.0 ...
```

Figura 5.2. Formati i skedarit *.asc* në eksperimentet e gjurmimit të syrit

5.4.1. Detektimi i ngjarjeve

Për detektimin dhe analizën e ngjarjeve të gjurmimit të syrit është përdorur libraria me burim të hapur `pymovements` (Krakowczyk, et al., 2023), e cila ofron një sërë algoritmesh për përpunimin dhe analizën e të dhënave të gjurmimit të syrit. Përpunimi është realizuar në mjedisin Python, duke u bazuar në konfigurimet paraprake, frekuencën e kampionimit dhe strukturën në formatin ASCII (*.asc*). Fillimisht është e rëndësishme të përcaktohet bashkësia e të dhënave që përmban të gjithë skedarët *.asc* të gjeneruar për secilin pjesëmarrës. Përmes funksioneve të librarisë `pymovements` kjo bashkësi të dhënash është ngarkuar në mjedisin e përpunimit për përllogaritje të mëtejshme.

Në fazën e parë të përpunimit, koordinatat e shikimit të regjistruara në njësi pixel janë konvertuar në gradë të këndit vizual *(degrees of visual angle)*, duke përdorur funksionin `gaze.pix2deg()` të librarisë `pymovements`. Ky transformim është i rëndësishëm për standardizimin e të dhënave dhe për aplikimin korrekt të shpejtësisë gjatë detektimit të ngjarjeve.

Më pas, pozicionet e shikimit janë shndërruar në sinjale shpejtësie përmes funksionit të `gaze.pos2vel()` duke përdorur një filtër Savitzky-Golay (Savitzky & Golay, 1964) për zbutjen e sinjalit pa humbur informacionin thelbësor. Duke ndjekur praktikat e ndjekura në literaturë të ngjashme (Salvucci & Goldberg, 2000), kemi përdorur një dritare prej 17 kampionesh për të dhënat me frekuencë kampionimi 1000 Hz, dhe një polinom të shkallës së dytë, me qëllim reduktimin e zhurmës pa humbur informacionin e rëndësishëm të sinjalit.

Zbulimi i fiksimeve është realizuar përmes një algoritmi të bazuar në shpejtësi (IVT-Velocity-Threshold Identification), sipas metodave të përshkruara në (Duchowksi,

2017) dhe (Salvucci & Goldberg, 2000). Një ngjarje është identifikuar si fiksim nëse shpejtësia e shikimit qëndron nën pragun maksimal prej 45 gradë për sekondë, për një kohëzgjatje minimale ndërmjet 60 dhe 100 ms (Duchowksi, 2017).

Ngjarjet të cilat nuk identifikohen si fiksime janë klasifikuar si sakada. Detektimi i sakadave është kryer duke ndjekur një prag minimal kohor prej 20 ms, sipas metodës së propozuar nga (Engbert, Trunkenbrod, Barthelme, & Wichman, 2015). Për këtë është përdorur funksioni i implementuar në librarinë `pymovements`:

```
gaze.detect('microsaccades',
            minimum_duration=20,
            threshold='engbert2015',
            name= 'saccade')
```

Funksioni i mësipërm me konfigurimet përkatëse lejon identifikimin e sakadave bazuar në ndryshimet e shpejtësisë së shikimit. Pas identifikimit të ngjarjeve është kryer përllogaritja e veçorive të tyre duke përfshirë metrikat e dispersionit të fiksimeve, amplitudën dhe shpejtësinë maksimale të sakadave, duke përdorur funksionin `gaze.event_properties()` të librarisë `pymovements`.

Më tej janë filtruar të dhënat duke ruajtur vetëm informacionin e rëndësishëm të gjurmimit të syrit nga skedarët ASCII. Informacioni i filtruar është ruajtur në file të tipit *.csv* të strukturuara për secilin pjesëmarrës. Këto file *.csv* janë përdorur si bazë për analizat e mëtejshme.

Në fazën përfundimtare janë gjeneruar zonat e interesit në nivel karakteri dhe fjale. Ky hap është realizuar përmes një skripti të posaçëm për gjenerimin e imazheve nga teksti i shfaqur në ekran, duke mundësuar përcaktimin e koordinatave hapësinore të secilit karakter. Ngjarjet e gjurmimit të syrit (fiksimet dhe sakadat) janë ndërlidhur me fjalët përkatëse duke përdorur shënuesin kohor (angl. *timestamp*) të fillimit dhe të përfundimit të leximit të secilit tekst, numrin identifikues të tekstit dhe numrin e faqes. Në këtë mënyrë, çdo ngjarje është lidhur me njësinë leksikore përkatëse duke mundësuar analizën e leximit në nivel fjale.

5.4.2. Përllogaritja e metrikave të leximit në nivel fjale

Duke u bazuar në ngjarjet e detektuara të gjurmimit të syrit, të përshkruara në seksionin 5.4.1, janë përllogaritur një sërë metrikash të leximit në nivel fjale, të cilat përdoren gjerësisht në studimet psikolinguistike për analizën e sjelljes gjatë leximit dhe të procesve të përpunimit gjuhësor. Këto metrika ofrojnë një bazë të pasur për analizën e sjelljes së leximit dhe studimin e mëtejshëm të kuptimit gjuhësor, specifikisht në rastin e punimit tonë për gjuhën shqipe.

Metrikat e llogaritura janë paraqitur në Tabelën 5.2, në të cilën përmblidhen përkufizimet e secilës prej metrikave. Një pjesë e metrikave lidhen me kalimin e parë të leximit, i cili reflekton proceset e para të identifikimit të fjalëve. Në këtë grupim përfshihet metrika e `first_fixation_duration` që përfaqëson kohëzgjatjen e fiksimit të parë mbi një fjalë dhe `first_duration` që përfaqëson kohëzgjatjen

totale të fiksimeve mbi një fjalë gjatë kalimit të parë. Të tjera metrika të llogaritura janë numri total i fiksimeve dhe rifiksimeve, metrika `skipped_first_pass` e cila identifikon fjalët që nuk janë fiksuar gjatë kalimit të parë.

Një grup tjetër i metrikave lidhet me kohën totale të leximit dhe rileximet (*re-reading*), që pasqyrojnë procese të përpunimit gjuhësor dhe semantik. Metrika `gaze_duration` përfaqëson kohën totale të leximit të një fjale gjatë kalimit të parë. Ndërkohë, metrika `total_fixation_duration` mat kohën e përgjithshme të kaluar duke lexuar një fjalë, duke përfshirë rikthimet. Metrika `go_past_time` mat kohën totale të kaluar nga momenti i fiksimit të parë deri në kalimin përtej një fjale, duke përfshirë rikthimet.

Tabela 5.2. Metrikat e përllogaritura nga leximi i teksteve në eksperimentet e gjurmimit të syrit

Emërtimi i metrikës	**Përshkrimi**
First duration	Kohëzgjata e fiksimeve gjatë kalimit të parë
First fixation duration	Kohëzgjatja e fiksimit të parë mbi një fjalë
Numri i rifiksimeve	Numri i rifiksimeve mbi një fjalë
Skipped in first pass	Numri i fjalëve të anashkaluara në kalimin e parë
Numri i fiksimeve	Kohëzgjatja e fiksimit të parë në një fjalë *w*
Gaze duration	Koha totale e leximit gjatë rikthimeve
Total fixation duration	Koha e leximit gjatë rikthimeve
Re-reading time	Koha e leximit gjatë rikthimeve
Second-pass duration	Koha e kaluar në kalimin e dytë
Go-past time	Koha deri në kalimin përtej fjalës, duke përfshirë rikthimet
In-word saccades	Numri i sakadave brenda fjalës
Incoming saccades	Numri i sakadave hyrëse drejt fjalës
Outcoming saccades	Numri i sakadave dalëse nga fjala

Metrikat si `in-word saccades`, `incoming saccades` dhe `outcoming saccades` shërbejnë si tregues të kuptimit të strategjive të leximit dhe vëmendjes vizuale dhe mënyrës se si një lexues përqendrohet në fjalë gjatë leximit. Këto metrika ofrojnë informacionin mbi mënyrën se si sytë lëvizin midis fjalëve dhe brenda tyre, duke reflektuar proceset konjitive që ndodhin gjatë përpunimit të tekstit. Analiza e tyre ndihmon në identifikimin e modeleve të përqendrimit dhe vështirësive të mundshme gjatë leximit.

5.4.3. Standardizimi i bashkësisë së të dhënave

Pas përllogaritjes së metrikave të leximit në nivel fjale për secilin pjesëmarrës, hapi i radhës ishte procesi i standardizimit dhe unifikimit të të dhënave për të krijuar një bashkësi të dhënash të përbashkët për studime të mëtejshme. Qëllimi kryesor i këtij hapi është njëtrajtësimi i të dhënave për të mundësuar analiza të mëtejshme gjuhësore, të cilat do të analizohen në KREU 6.

Hapi i parë i standardizimit të bashkësisë së të dhënave ishte bashkimi i të dhënave të të gjithë pjesëmarrësve. Çelësat të cilat u përdorën për bashkimin në një bashkësi të vetme janë kolonat identifikuese të cilat ruajnë përputhshmërinë midis tekstit, faqes, pozicionit të fjalës dhe vetë fjalës:

```
GROUP_COLS =
["text_id", "page_number", "word_number", "word"]
```

Për secilën fjalë u bashkuan metrikat kryesore të leximit, si më poshtë:

```
NUM_COLS = [
    "first_duration",
    "first_fixation_duration",
    "gaze_duration",
    "total_fixation_duration",
    "re_reading_time",
    "second_pass_duration",
    "go_past_time",
    "number_of_fixations",
    "number_of_inword_saccades",
    "number_of_outcoming_saccades",
    "number_of_incoming_saccades"
]
```

Për secilën metrikë u përllogaritën statistika përmbledhëse standarde, konkretisht: *mesatarja, mediani, devijimi standard, gabimi standard i mesatares, vlerat minimale, maksimale* dhe *numri total i vëzhgimeve* ndër pjesëmarrësit. Ky bashkim dhe këto fusha përmbledhëse mundësojnë një përfaqësim të sjelljes gjatë leximit në nivel fjale ndër pjesëmarrësit e përfshirë në studim.

```
for col in NUM_COLS:
    AGG[col] = [
        "mean",
        "median",
        "std",
        "sem",
        "min",
        "max",
        "count"
    ]
```

Në vijim, bashkësia e të dhënave e krijuar pas hapave të sapo përmendura, u përshtat me strukturën e studimeve gjuhësore. Konkretisht, është realizuar ndarja e njësive shumëfjalëshe (angl. *multiword token*), bazuar në informacion e marrë nga formati CoNLL-u i etiketuar manualisht nga ekspertë të fushës për secilin tekst. Ky hap është i domosdoshëm, meqenëse një nga qëllimet e përdorimit të të dhënave për gjurmimin e syrit është ndërthurja me etiketimet gjuhësore.

Përveç kësaj, u trajtuan shenjat e pikësimit të bashkëngjitura në fillim ose në fund të fjalëve. Në rastin e eksperimenteve të gjurmimit të syrit, zona e interesit ndahen sipas hapësirave, duke sjellë bashkëngjitjen e shenjave të pikësimit me fjalët. Për të krijuar lidhjen një më një me secilin token që ndodhet në etiketimet manuale gjuhësore, shenjat e pikësimit u identifikuan dhe u ndanë nga fjala përkatëse dhe metrikat e leximit u shpërndanë në mënyrë proporcionale, për të shmangur rritjen artificiale të kohëzgjatjeve të fiksimit drejt shenjave të pikësimit.

Përfundimisht, bashkësia e të dhënave është normalizuar dhe standardizuar, në mënyrë që të eliminohej efekti i shkallëve të ndryshme matëse dhe për të arritur një stabilitet më të madh të modeleve të ndryshme të mësimit të makinës që do të përdoren në vijim. Si rezultat u gjenerua një bashkësi të dhënash e standardizuar në nivel fjale, që shërben për analizat e mëtejshme dhe për integrimin me etiketimet gjuhësore, veçoritë morfologjike dhe varësitë sintaksore.

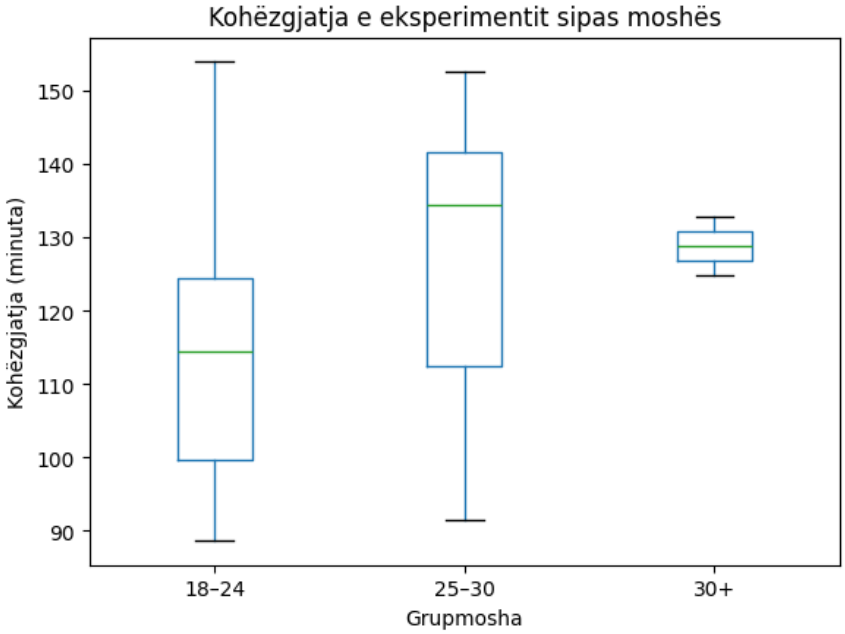
Theksojmë se të gjithë hapat e përdorur mund të replikohen dhe ndryshohen hap pas hapi në rastin se nevojiten të kryhen studime të natyrave të ndryshme.

5.5. Analiza e sjelljes gjatë leximit të pjesëmarrësve

Për të kuptuar sjelljen e leximit të pjesëmarrësve në eksperimentet e zhvilluara, kemi përdorur të dhënat e grumbulluara gjatë eksperimentit, të dhënat e mbledhura gjatë pyetësorit të pjesëmarrësve dhe përgjigjet e dhëna nga përdoruesi ndaj pyetjeve kuptimore pas secilit tekst. Në vijim do të prezantojmë një analizë të detajuar e cila përqendrohet në marrëdhëniet ndërmjet faktorëve demografikë (mosha, gjinia, niveli i arsimit) dhe gjuhëve të ndryshme në të cilat pjesëmarrësi ka njohuri.

Analiza e parë e kryer përqendrohet në vlerësimin e kohëzgjatjes në minuta të eksperimentit për secilin nga pjesëmarrësit. Pjesëmarrësit janë ndarë në 3 grupmosha: 18-24 (11 pjesëmarrës), 25-30 vjeç (10 pjesëmarrës) dhe mbi 30 vjeç (2 pjesëmarrës).

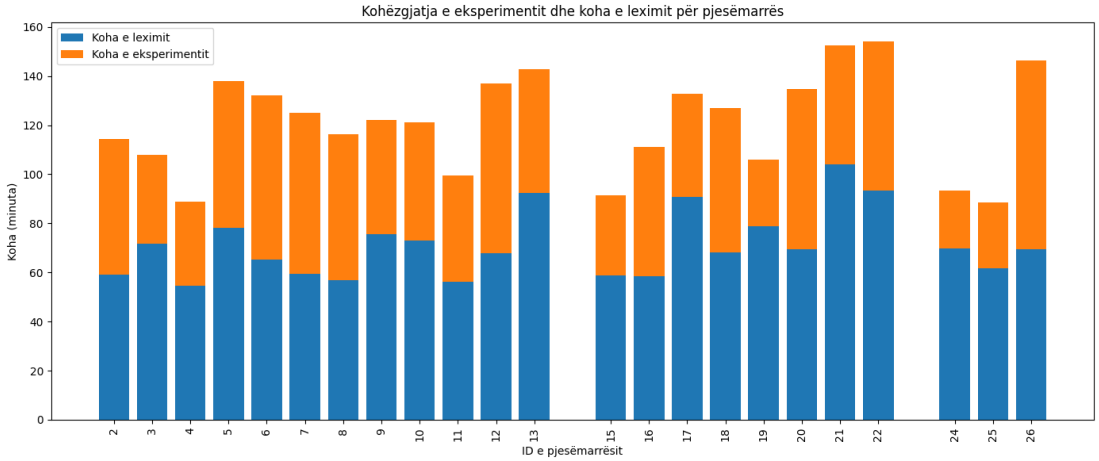
Duke qenë se pyetësori demografik nuk ishte i detyruar për t'u kryer, dy nga pjesëmarrësit nuk e kanë plotësuar atë, prandaj të dhënat e paraqitura do të jenë për 23 pjesëmarrësit e tjerë (duke mos përfshirë pjesëmarrësit me id 14 dhe 23). Në Grafikun 5.1 mund të shikohet kohëzgjatja e eksperimentit në grupmosha të ndryshme, ku vihet re se moshat më të reja kanë një kohëzgjatje mesatare të zhvillimit të eksperimentit më të vogël se moshat e tjera.



Grafiku 5.1. Kohëzgjatja e eksperimentit të gjurmimit të syrit sipas moshës

Në Grafikun 5.2 paraqitet diferenca midis kohëzgjatjes së eksperimentit dhe kohës totale që është shpenzuar për të lexuar tekstet. Në kohëzgjatjen e gjithë eksperimentit, shënuar në ngjyrë portokalli, përfshihet koha nga fillimi e deri në përfundim të eksperimentit të leximit, duke përfshirë kohën e leximit, kohën për kalibrimet, validimet, pushimet mes teksteve dhe kohën e shpenzuar për t'u përgjigjur pyetjeve kuptimore. Nga ky grafik shikojmë se kemi një shpërndarje relativisht të qëndrueshme ndërmjet pjesëmarrësve të ndryshëm.

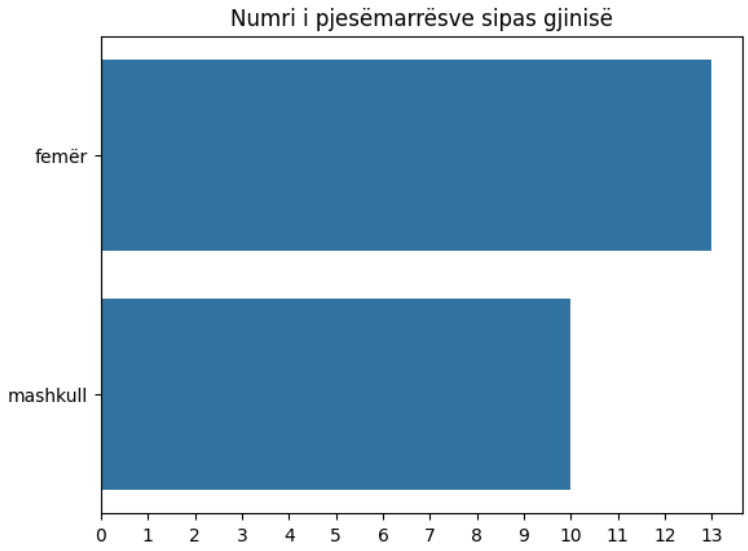
Mesatarja e kohës totale të leximit është 71.01 minuta, me një devijim standard prej 13.46 minutash, që paraqet një variabilitet të moderuar në shpejtësinë e leximit. Pjesëmarrësi me vlerën më të afërt me mesataren e kohës së leximit është ai me numër identifikues 3, me një kohë leximi prej 71.56 minuta, dhe 16 pjesëmarrës të tjerë rezultojnë me vlera afër mesatares, që tregon një ritëm të ngjashëm leximi gjatë eksperimentit. Devijimet më të mëdha i shikojmë në pjesëmarrësin me numër identifikues 21, me një kohë prej 104.09 minuta. Faktorë të ndryshëm që mund të ndikojnë në këto devijime janë strategjia e leximit, niveli i përqendrimit, lodhja dhe dallimet në kompetencat gjuhësore, gjë që thekson rëndësinë e ndarjes së pjesëmarrësve sipas këtyre veçorive për të studiuar korrelacionin midis faktorëve të ndryshëm dhe kohës së leximit.



Mesatarja	71.01 minuta	
Deviacioni standard	13.46 minuta	
Vlera më afër mesatares	*ID e pjesëmarrësit*	3
	Koha e leximit	71.57 minuta
	Ndryshesa nga mesatarja	0.54 minuta
Vlera më e largët nga mesatarja	*ID e pjesëmarrësit*	21
	Koha e leximit	104.08 minuta
	Ndryshesa nga mesatarja	33.07 minuta
Numri i pjesëmarrësve afër kohës mesatare	16	
Numri i pjesëmarrësve larg kohës mesatare	7	

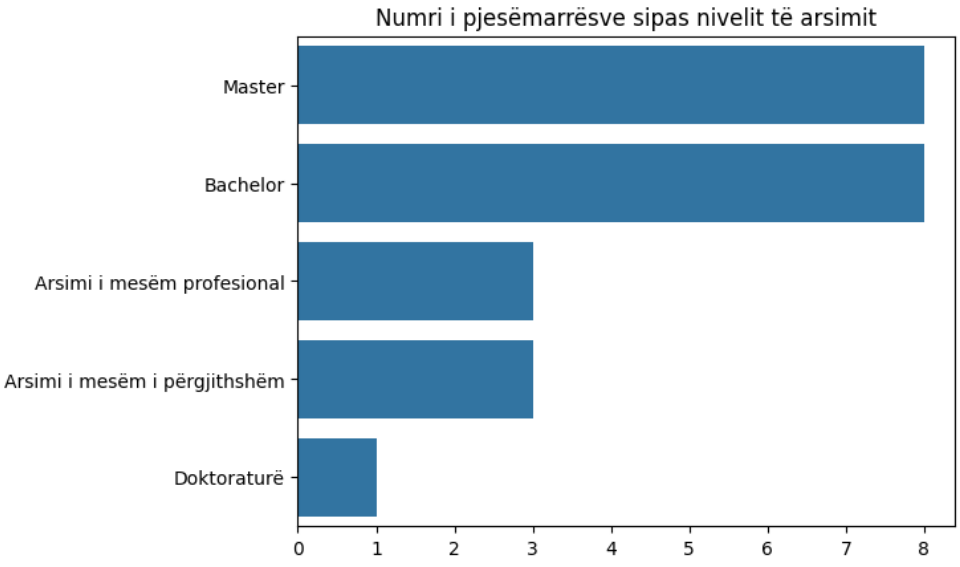
Grafiku 5.2. Kohëzgjatja e eksperimentit të gjurmimit të syrit sipas pjesëmarrësve: mesatarja dhe ekstremet

Në Grafikun 5.3 paraqitet shpërndarja e pjesëmarrësve sipas gjinisë, ku vlerësohet edhe një balancë gjinore midis pjesëmarrësve në eksperiment, duke sjellë që rezultatet të mos jenë jo proporcionale sipas një gjinie të caktuar. Kjo balancë kontribuon në rritjen e besueshmërisë dhe përfaqësueshmërisë së të dhënave të mbledhura gjatë studimit.



Grafiku 5.3. Shpërndarja e pjesëmarrësve në eksperimentin e gjurmimit të syrit sipas gjinisë

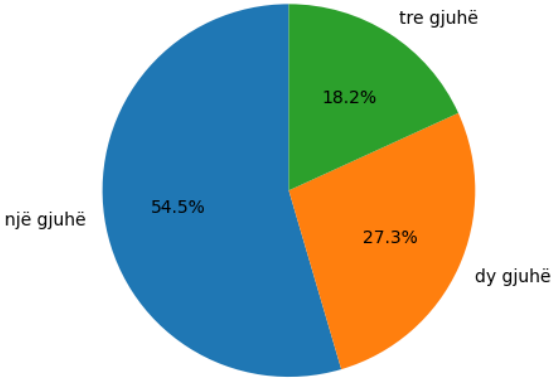
Ndërkohë në Grafikun 5.4 numri i pjesëmarrësve sipas nivele të ndryshme të arsimit tregon që shumica e pjesëmarrësve (17 prej tyre) kanë arsim të lartë (Bachelor, Master, Doktoraturë) dhe vetëm 6 prej tyre janë me arsim të mesëm të përgjithshëm ose profesional. Kjo shpërndarje tregon se kampioni i studimit përbëhet kryesisht nga individë me nivel më të avancuar arsimor, që mund të ndikojë edhe në strategjitë e leximit dhe në mënyrën e përpunimit të tekstit gjatë eksperimentit.



Grafiku 5.4. Shpërndarja e pjesëmarrësve në eksperimentin e gjurmimit të syrit sipas nivelit të arsimit

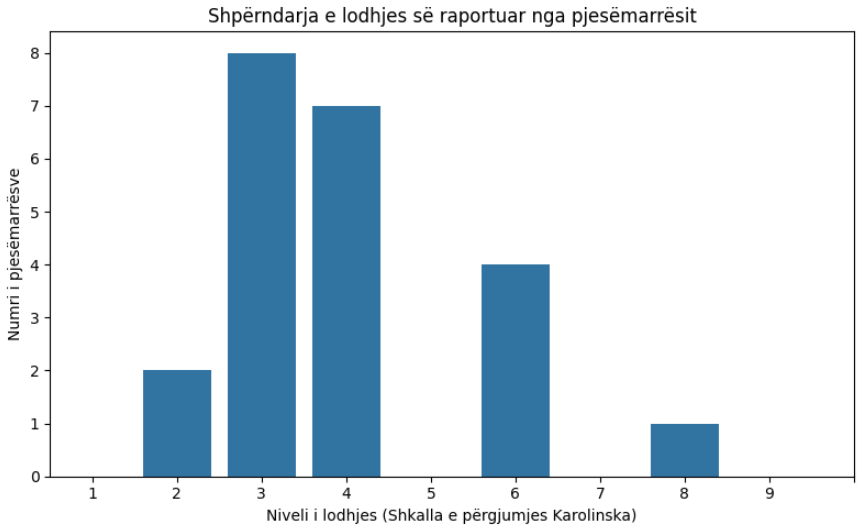
Grafiku 5.5 paraqet larmishmërinë e kompetencave gjuhësore të pjesëmarrësve në eksperiment. Shikojmë se 45.5% e pjesëmarrësve janë rritur me dy ose tre gjuhë gjatë fëmijërisë. Kjo përqindje e pjesëmarrësve përfaqëson pjesëmarrësit nga diaspora, duke

qenë se janë rritur në Zyrih dhe në fëmijëri kanë përdorur disa gjuhë në komunikimin e përditshëm. Duke qenë se një pjesë prej tyre, gjuhën shqipe e kanë përvetësuar vetëm nga komunikimi nga familja, në sjelljet e tyre gjatë leximit vihet re një ngadalësi më e madhe për të kuptuar tekstet e shkruara në gjuhën shqipe. Kjo situatë mund të ndikojë edhe në metrikat e leximit të përllogaritura, duke reflektuar dallime në procesin e përpunimit të tekstit.



Grafiku 5.5. Shpërndarja e pjesëmarrësve në eksperimentin e gjurmimit të syrit sipas gjuhës së fëmijërisë

Një tjetër pyetje e rëndësishme e përfshirë në pyetësorin demografik është edhe përcaktimi i shkallës së përgjumjes sipas 'Shkallës së Përgjumjes Karolinska' (Shahid, Wilkinson, Marcu, & Shapiro), paraqitur në Grafikun 5.6. Kjo shkallë përcakton një njësi matëse për nivelin e përgjumjes ose lodhjes së pjesëmarrësit në një kohë të caktuar. Kjo është relativisht e rëndësishme për të kuptuar gjendjen e pjesëmarrësit gjatë eksperimentit.



Grafiku 5.6. Shpërndarja e pjesëmarrësve në eksperimentin e gjurmimit të syrit sipas nivelit të lodhjes (shkalla e përgjumjes Karolinska)

Në rastin e studimit tonë është përdorur një shkallë me 9 gradë (1 - jashtëzakonisht i vëmendshëm, 2 - shumë i vëmendshëm, 3 - i vëmendshëm, 4 - mesatarisht i vëmendshëm, 5 - as i vëmendshëm, as i përgjumur, 6 - me disa shenja përgjumje, 7 - i përgjumur, por pa përpjekje për të qëndruar i vëmendshëm, 8 - I përgjumur, pak përpjekje për të qëndruar i vëmendshëm, 9 - shumë i përgjumur, me shumë përpjekje për të qëndruar i vëmendshëm, duke luftuar me përgjumjen).

Nga Grafiku 5.6 vërehet se shumica e pjesëmarrësve janë kryesisht të vëmendshëm, me një vlerë më të vogël se 5 në shkallën Karolinska, ndërkohë 5 pjesëmarrës tregojnë disa shenja lodhjeje apo përgjumjeje. Kjo tregon se në përgjithësi pjesëmarrësit kanë qenë në një gjendje të përshtatshme për realizimin e eksperimentit dhe për përqendrim gjatë leximit të teksteve.

5.6. Krahasimi i metrikave të leximit në nivel fjale mes eksperimenteve të gjurmimit të syrit dhe mausit

Pas krijimit të dy bashkësive të të dhënave, njëra nga eksperimentet e gjurmimit të syrit dhe tjetra nga eksperimentet e gjurmimit të mausit, kemi analizuar sa përfaqësuese janë të dhënat nga gjurmimi i mausit dhe sa mirë reflektojnë ato procesin e lëvizjes së syrit. Për këtë arsye kemi kryer një vlerësim praktik të implikimeve që mund të ndodhin gjatë përdorimit të metodave me kosto të ulët (si p.sh. gjurmimi i mausit) në studimet e lidhura me leximin.

Qëllimi kryesor është të krahasohen metrikat e leximit nga eksperimentet e gjurmimit të syrit dhe mausit. Më pas të vlerësohet korrelacioni në nivel fjale duke përdorur analizën e korrelacionit Bayesian (angl. *Bayesian correlation analysis*). Në këtë mënyrë vlerësojmë limitimet e mundshme të metodës së gjurmimit të mausit.

Për të përgatitur secilën bashkësi të të dhënave, të gjithë skedarët e pjesëmarrësve janë bashkuar në një *dataframe* të vetme. Në këtë mënyrë kemi krijuar dy *dataframe* të ndërlidhur me secilën metodë eksperimentale. Hapi i rradhës ishte mesatarizimi i vlerave në nivel teksti-faqe-fjale. Duke përdorur indeksimin në nivel fjale në secilin tekst, dy bashkësitë e të dhënave u bashkuan në një *dataframe* të vetëm, ku për secilën fjalë kishim metrikat si: `first duration`, `total fixation duration`, `gaze duration` dhe `go-past time` (Haveriku, Kote, & Meçe, 2025).

Për të vlerësuar lidhjen mes metrikave nga gjurmimi i mausit dhe i syrit aplikuam korrelacionin Bayesian, bazuar në punime të ngjashme të kryera për gjuhën angleze (Wilcox, Ding, Sachan, & Jäger, 2024). Kjo metodikë mundëson informacionin e nevojshëm duke llogaritur shpërndarjen posteriore dhe koeficientin e korrelacionit. Analiza u zhvillua në mënyrë të veçantë për secilën nga metrikat e leximit të studiuara.

Kampionimi posterior u krye përmes metodës Markov Chain Monte Calo (MCMC), duke përdorur 2000 hapa përmirësimi (angl. *tuning*) dhe duke përllogaritur për secilën metrikë mesataren posteriore (angl. *posterior mean*) dhe intervalin e densitetit të lartë 95% (angl. *highest density interval*) (Behseta, Berdyyeva, Olson, & Kass, 2009).

Në Tabelën 5.3 përmblidhen të gjitha rezultatet e arritura për secilin tekst dhe secilën metrikë leximi. Rezultatet tregojnë një korrelacion pozitiv, duke sugjeruar që gjurmimi

i mausit arrin të kapë në mënyrë të mirë mënyrën e leximit (angl. *reading-time patterns*). Ndërkohë në Figurën 5.3 mund të vizualizojmë vlerat për secilin tekst.

Tabela 5.3. Korrelacioni Bayesian midis metrikave të leximit nga eksperimentet e gjurmimit të syrit dhe të mausit

		Total Duration	**First Duration**	**Gaze Duration**	**Go-past Time**
Teksti 1: Hëna	*Posterior mean ρ 95% CI*	0.621 [0.573, 0.672]	0.539 [0.481, 0.596]	0.314 [0.235, 0.395]	0.490 [0.422, 0.551]
Teksti 2: Prilli i Thyer	*Posterior mean ρ 95% CI*	0.621 [0.57, 0.67]	0.541 [0.482, 0.599]	0.314 [0.230, 0.391]	0.492 [0.429, 0.553]
Teksti 3: Njeriu i Shpellave	*Posterior mean ρ 95% CI*	0.401 [0.339, 0.459]	0.331 [0.266, 0.400]	0.191 [0.113, 0.264]	0.146 [0.072, 0.224]

Kohëzgjatja totale pasqyron shumën e të gjithë fiksimeve në një fjalë. Në rezultatet e paraqitura në Tabelën 5.3, vërejmë se metrika e kohëzgjatjes totale shfaq korrelacionin më të fortë, duke variuar nga 0.401 deri në 0.621. Metrika e kohëzgjatjes së parë (angl. *first duration*) pasqyron kohëzgjatjen e fiksimit të parë mbi një fjalë, përpara se të ndodhin regresione. Nga Tabela 5.3 vërejmë se korrelacioni i kësaj metrike është i lartë në rastin e tekstit të parë dhe të dytë, por zvogëlohet në tekstin e tretë. Ky rezultat sugjeron se të dhënat nga eksperimentet e gjurmimit të mausit arrijnë të përafrojnë mirë në rastin e teksteve të thjeshta, por kanë vështirësi në tekstet më komplekse apo shkencore.

Metrika e tretë, e kohëzgjatjes së vështrimit (angl. *gaze duration*) pasqyron shumën e të gjithë fiksimeve mbi një fjalë gjatë kalimit fillestar. Rezultatet e paraqitura në Tabelën 5.3 tregojnë se kjo metrikë shfaq korrelacione të larta për tekstin e parë dhe të dytë, por bie për rastin e tekstit të tretë. Vlera të ngjashme vërehen edhe për metrikën e kohës së kthimit pas (angl. *go-past time*). Kjo rënie mund të lidhet me krakateristika specifike të tekstit të tretë, si struktura sintaksore ose niveli i vështirësisë së fjalorit. Një ndryshim i tillë sugjeron se përpunimi i këtij teksti mund të ketë kërkuar strategji të ndryshme leximi nga pjesëmarrësit, duke mos u lidhur drejtpërdrejt me metodikën e përdorur për zhvillimin e eksperimenteve.

Si përfundim rezultatet e arritura, tregojnë një korrelacion të mirë në rastin e teksteve 'Hëna' dhe 'Prilli i Thyer', veçanërisht për metrikën e kohëzgjatjes totale. Në kontrast, në tekstin e tretë 'Njeriu i Shpellave' rezultatet tregojnë një korrelacion më të dobët për të gjitha metrikat, duke sugjeruar se teknika e gjurmimit të mausit është më e ndjeshme ndaj karakteristikave të tekstit, të cilat ndikojnë drejtpërdrejt në ndërveprimin e lexuesit.

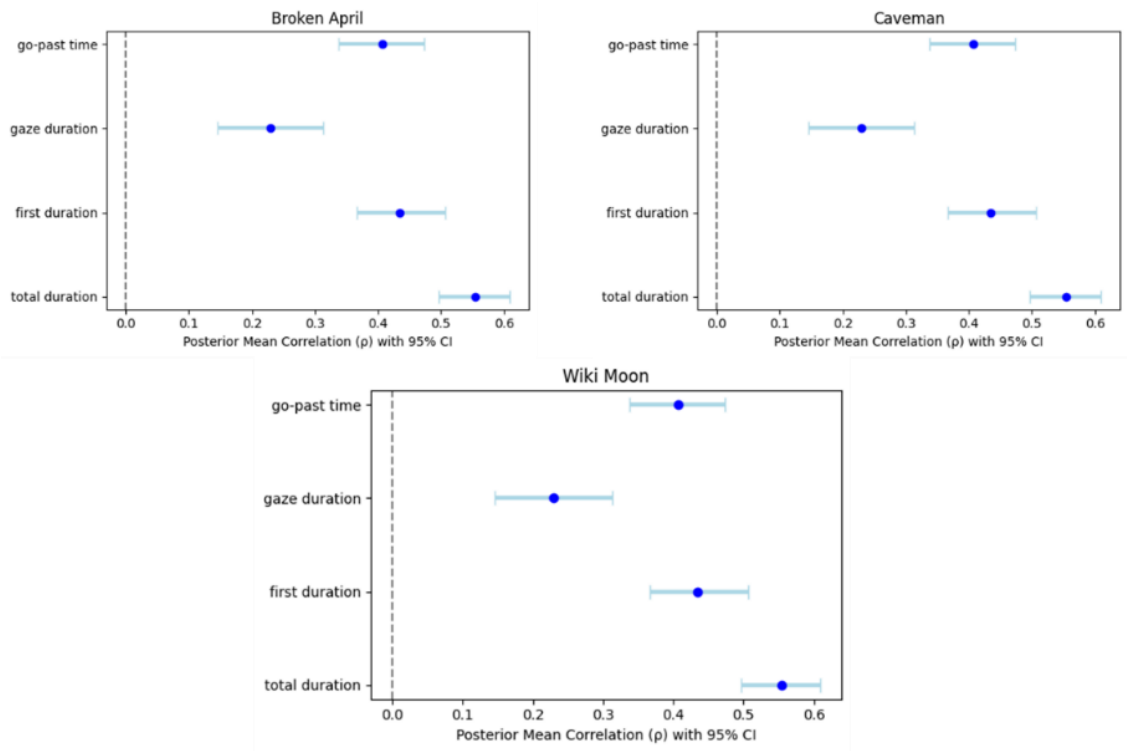


Figura 5.3. Korrelacioni ndërmjet metrikave të leximit nga gjurmimi i syrit dhe i mausit

Megjithatë, rezultatet e përgjithshme tregojnë se teknika e gjurmimit të mausit mund të shërbejë si një alternative e realizueshme me kosto të ulëta për të përafruar sjelljen gjatë leximit. Kërkimet në të ardhmen mund t'i zgjerojnë këto rezultate duke përfshirë dhe studiuar metrika të tjera të leximit, si edhe duke përfshirë tekste më të larmishme.

Gjithashtu, rritja e numrit të pjesëmarrësve në secilën nga bashkësitë e të dhënave mund të ndihmojë në reduktimin e zhurmës dhe në përmirësimin e përgjithësimit të rezultateve. Për më tepër, të dhënat nga gjurmimi i mausit mund të kombinohen me burime të tjera të të dhënave konjitive, duke ofruar një kuptim më të plotë tëproceseve të leximit. Kjo do të mundësonte zhvillimin e modeleve më të sakta për përpunimin e teksteve dhe analizën e sjelljes së lexuesve.

KREU 6

INTEGRIMI I TË DHËNAVE KONJITIVE NË MODELET E PGJN-së

6.1. Hyrje

Integrimi i të dhënave konjitive, nga eksperimentet e gjurmimit të mausit apo syrit, për të përmirësuar modelet ekzistuese të PGJN-së, në punimet për gjuhët më të përhapura kanë arritur në rezultate të mira të cilat motivojnë punimet e mëtejshme në këtë fushë. Shumë studime, të listuara dhe analizuara në nënkapitullin 3.2, kanë theksuar që përfshirja e informacionit nga leximi i teksteve gjatë fazave të trajnimit apo testimit përmirëson saktësinë e modeleve të mësuarit makinë ose të rrjeteve nervore, në fusha si analiza e ndjenjës, etiketimi i pjesëve të ligjëratës dhe identifikimi i fjalëve kyçe.

Një sfidë e identifikuar nga rishikimi i literaturës në kreun 3 ishte nevoja e zgjerimit të të dhënave konjitive nga proceset e përpunimit të gjuhës gjatë leximit në shumë gjuhë, sidomos në gjuhët me burime të pakta. Zhvillimi i eksperimenteve në gjuhë më pak të studiuara mund të ndihmojë në zbulimin e ngjashmërive dhe ndryshimeve në mënyrën si njerëzit përpunojnë informacionin në gjuhë të ndryshme. Për më tepër theksuam se kërkimet e shumta kanë provuar që përdorimi i mjeteve me kosto të ulët si kamerat web, pajisjet mobile apo pajisja maus, janë mjete të cilat ofrojnë saktësi të pranueshme në mungesë të laboratorëve të mirëfillta.

Potenciali i të dhënave nga eksperimentet e gjurmimit të syrit, nevoja e studimeve paralele shumëgjuhësore dhe përdorimi i pajisjeve alternative, u prezantuan teorikisht në kreun 3. Ndërkohë, në kreun 4 u paraqit metodologjia e ndjekur nga eksperimentet e gjurmimit të mausit dhe në kreun 5 metodologjia për eksperimentet e gjurmimit të syrit. Bashkësitë e të dhënave të përshkruara në kreun 4 dhe kreun 5 do të përdoren praktikisht në këtë kapitull për të eksperimentuar me përdorimin e të dhënave konjitive për gjuhën shqipe.

Pyetjet kryesore kërkimore të cilat motivojnë punën e përshkruar në këtë kapitull janë:

1. A përmirëson gjurmimi i syrit apo i mausit performancën e modeleve të etiketimit automatik të pjesëve të ligjëratës dhe lidhjeve sintaksore?
2. Cilat janë kategoritë gramatikore të cilat përfitojnë më shumë nga sinjalet konjitive?
3. A përmirësohet identifikimi i fjalëve kyçe në tekst përmes përdorimit të sinjaleve konjitive nga gjurmimi i syrit/ mausit?
4. Cilat janë metrikat e leximit më informuese për secilën problematikë?
5. Si ndryshon ndikimi i të dhënave konjitive sipas detyrës (sintaksore apo semantike)?

Në këtë kapitull prezantohet metodologjia e përdorimit e bashkësisë së të dhënave të prezantuara në kreun 4 dhe kreun 5 dhe integrimi i tyre në modelet vetëm me veçori gjuhësore të inteligjencës artificiale.

6.2. Konteksti aktual për gjuhën shqipe

Gjuha shqipe është pjesë e një dege të veçantë në familjen e gjuhëve indoevropiane me rreth 8 milionë folës. Ajo ndahet në dy dialekte kryesore, dialekti gegë, i folur në veriun e Shqipërisë, në Kosovë, në komunitetet shqiptare në Maqedoninë e Veriut dhe në Malin e Zi, dhe dialekti toskë, i folur kryesisht në jug të Shqipërisë, duke përfshirë edhe diasporën arbëreshe në jug të Italisë dhe arvanitase në Greqi (Hyllested & Joseph, 2022). Gjuha standarde shqipe bazohet kryesisht në dialektin toskë. Kërkimet në fushën e PGJN-së për gjuhën shqipe, si një gjuhë me burime të pakta, kanë ndeshur shumë vështirësi. Ndër vite janë bërë përpjekje të shumta, por të fragmentuara, për të studiuar aspekte të ndryshme të PGJN-së për gjuhën shqipe, por shpesh këto nisma kanë qenë të kufizuara në përmasa, me burim të mbyllur ose kanë pasur mungesë të përfshirjes të ekspertëve gjuhësorë për studimin e çështjeve të ndryshme (Haveriku & Frashëri, 2023), (Haveriku & Meçe, 2025).

Ndërkohë fusha e gjurmimit të syrit nuk është shumë e lëvruar për studimin e sjelljeve gjatë leximit për gjuhën shqipe. Duke u bazuar në punimet e përmbledhura në nënkapitullin 3.2, përdorimi i të dhënave konjitive ka treguar përmirësime të vlefshme në nënfusha të ndryshme të PGJN-së. Të dhënat konjitive mund të ndihmojnë në lehtësimin dhe shpejtimin e procesit të etiketimit gjuhësor, i cili zakonisht kërkon kohë të konsiderueshme dhe ekspertizë të specializuar gjuhësore (Barret, Keller, & Sogaard, 2016). Duke marrë në konsideratë kohën dhe ekspertizën e kërkuar për të etiketuar korpuse të mëdha, të dhënat nga leximi ofrojnë një alternativë të shpejtë për procesin e etiketimit. Gjithashtu, pjesëmarrësit në eksperimente të gjurmimit të syrit apo të mausit nuk duhet të jenë ekspertë të fushës, mjafton që të jenë folës amtarë të gjuhës shqipe. Kjo e lehtëson akoma më tej grumbullimin e të dhënave, pasi kriteret e përfshirjes së pjesëmarrësve në eksperiment janë relativisht minimale.

Dy nga korpuset më të mëdha që ekzistojnë për gjuhën shqipe janë korpusi i gjuhës shqipe i zhvilluar nga Instituti i Shën Petersburgut (Morozova, Rusakov, & Arkhangelskij, 2025) dhe korpusi sqGlove i krijuar nga Universiteti i Studimeve të Gjuhëve të Huaja në Beijing (Ke, et al., 2012). I pari është i përbërë nga 16.6 milion token, ndërsa i dyti nga 1 milion fjalë, secili i etiketuar me etiketat e pjesëve të ligjëratës dhe temat e fjalëve. Megjithatë, të dyja këto punime, janë me qasje të kufizuar duke penguar mundësinë e aksesueshmërisë dhe zhvillimit të punimeve të mëtejshme.

Në kuadër të kornizës (angl. *framework*) Universal Dependencies (UD) (Nivre, et al., 2020) aktualisht ekzistojnë tre treebank-e të disponueshme për gjuhën shqipe. I pari është TSA (Treebank of Standard Albanian), i cili përbëhet nga 60 fjali, 922 token, të etiketuar me pjesët e ligjëratës dhe lidhjet sintaksore (Toska, Nivre, & Zeman, 2020). I dyti, STAF (Saarbruecken Treebank of Albanian Fiction), përbëhet nga 202 fjali dhe 2499 tokens (Talamo, 2025). I treti, GPS (Gheg Pear Stories) është një treebank i dedikuar dialektit Geg, i përbërë nga 966 fjali, 15990 token (Ebert, et al., 2022).

Një tjetër korpus në zhvillim e sipër për gjuhën shqipe sipas standardeve të kornizës së UD është edhe SALT (The Standard Albanian Language Treebank) i cili përbëhet nga 1400 fjali dhe 24537 token. Kjo bashkësi e të dhënave përfshin segmentimin e fjalive dhe fjalëve, etiketimet e pjesëve të ligjëratës, lematizimin, veçoritë morfologjike dhe veçoritë sintaksore (Kote, et al., 2024).

Ndërkohë, puna kërkimore për mbledhjen dhe përpunimin e të dhënave konjitive nga gjurmimi i mausit dhe i syrit për gjuhën shqipe është ende në faza fillestare. Në (Haveriku, Haxhija, & Meçe, 2024) autorët prezantojnë një korpus të vogël të bazuar në gjurmimin e syrit përmes kamerës së kompjuterit nga 10 pjesëmarrës. Ky punim bazohet në përdorimin e librarisë WebGazer (Papoutsaki, et al., 2016). Përveç kësaj, një korpus i krijuar nga gjurmimi i kursorit të mausit (Haveriku, Bedulla, & Meçe, 2025), duke përshtatur mjetin MoTR (Wilcox, Ding, Sachan, & Jäger, 2024) për gjuhën shqipe, është një tjetër kontribut në këtë fushë për gjuhën shqipe. Ky korpus është krijuar nga zhvillimi i eksperimentit të gjurmimit të mausit me 51 folës amtarë të gjuhës shqipe dhe është përdorur për të zhvilluar eksperimente me integrimin e bashkësisë së të dhënave nga gjurmimi i mausit për parashikimin e etiketave të pjesëve të ligjëratës dhe lidhjeve sintaksore (Haveriku, et al., 20225).

6.2.1. Standard Albanian Language Treebank (SALT)

SALT (Standard Albanian Language Treebank) është një korpus sintaksor i shqipes standarde i cili përbëhet nga 27213 fjalë, 1556 fjali dhe përfshin etiketimet e lidhjeve sintaksore, etiketat e pjesëve të ligjëratës, veçoritë morfologjike dhe temën (Kote, et al., 2024). Ky korpus është ndërtuar duke u bazuar në rregullat dhe formatin e kornizës së Universal Dependencies (UD). Gjatë procesit të etiketimit, ekspertët gjuhësorë kanë treguar ruajtjen e pasurisë morfologjike të shqipes standarde, duke përshtatur bashkësinë e etiketave të ofruara nga UD dhe duke adresuar veçoritë specifike të lidhura me gjuhën shqipe.

Në Tabelën 6.1 paraqiten disa statistika kryesore të lidhura me korpusin SALT dhe frekuencën e shfaqjes së etiketave kryesore në këtë bashkësi të dhënash (Kote, et al., 2024). Të gjitha fjalitë e përfshira në korpus janë përzgjedhur nga burime të hapura, libra artistikë, tekste gramatikore dhe korpusi Leipzig (Golhahn, Eckart, & Quasthoff, 2012). Para nisjes së procesit të etiketimit fjalitë janë kontrolluar nga ekspertë gjuhësorë për gabime të ndryshme gramatikore, si mungesa e shkronjave si “ë”, “ç”, gabime tipografike etj. Për të lehtësuar procesin e etiketimit, fillimisht është bërë etiketimi automatik i fjalive duke përdorur modelin e propozuar nga Kote et al. (2019) për segmentimin, lematizimin, etiketimin e pjesëve të ligjëratës dhe veçorive morfologjike.

Fjalitë e etiketuara automatikisht janë kontrolluar dhe rregulluar në rast të gabimeve të mundshme, e më pas i është shtuar një nivel tjetër etiketimi, ai sintaksor, i cili është kryer në mënyrë manuale nga ekspertët gjuhësorë. Për përcaktimin e saktë të temës së fjalës (angl. *lemma*) ekspertët e gjuhësisë kanë përdorur *Fjalorin e gjuhës së sotme shqipe* (Akademia e Shkencave e Shqipërisë, 1980, 2002, 2006).

Tabela 6.1. Statistika mbi korpusin SALT

Numri i fjalive		1,556 fjali	
Numri i fjalëve		27,213	
Numri i *shprehjeve shumëfjalëshe* (angl. *multiword tokens*)		87	
Gjatësia mesatare e fjalive		17.38	
Frekuenca e etiketave			
UPOS	**Frekuenca**	**Deprel**	**Frekuenca**
NOUN	5,845	punct	3,246
PUNCT	3,254	det	2,504
DET	2,969	case	2,279
VERB	2,686	advmod	2,042
ADP	2,318	nsubj	1,803
PRON	2,100	nmod	1,561

Ndërkohë gjatë procesit të etiketimit manual janë marrë në konsideratë veçoritë leksiko-gramatikore sipas Gramatikës së gjuhës shqipe të publikuar nga Akademia e Shkencave e Shqipërisë (Agalliu, et al., 2002). Në përfundim janë përdorur 17 etiketa të pjesëve të ligjëratës dhe 48 etiketa të marrëdhënieve sintaksore nga korniza e UD.

Dy aplikacione të cilat janë përzgjedhur nga grupi teknik i projektit për lehtësimin e procesit të etiketimit nga ekspertët gjuhësorë janë Conllu Editor (Heinecke, 2019) dhe Arborator Grew (Guibon, Courtin, Gerdes, & Guillaume, 2020), të paraqitur me shembull në Figurën 6.1.

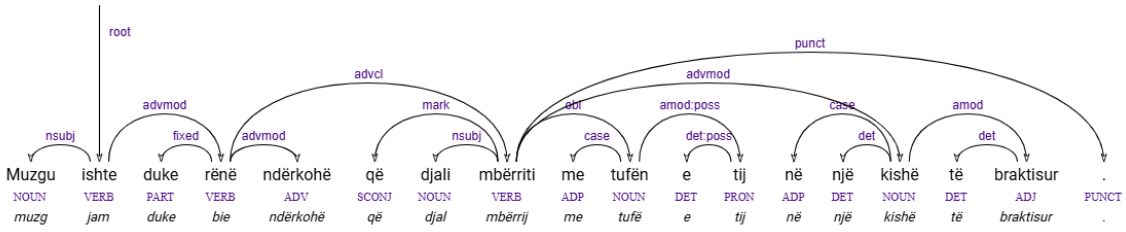


Figura 6.1. Paraqitje vizuale e etiketimit të pjesëve të ligjëratës dhe lidhjeve sintaksore në platformën Arborator Grew

Korpusi SALT është në fazat përfundimtare të përpunimit dhe pastrimit, para se të publikohet i plotë për gjithë bashkësinë e kërkuesve të interesuar në këtë fushë. Ky korpus është përdorur si bazë eksperimentale për të gjitha eksperimentet e prezantuara në pjesën në vijim. Ai përfshin tekste të përpunuara dhe të etiketuara në nivele të ndryshme gjuhësore, duke mundësuar zhvillimin dhe testimin e metodave të ndryshme të përpunimit të gjuhës natyrore.

6.2.2. Korpuse për gjuhën shqipe në kornizën UD

6.2.2.1. UD Treebank for Standard Albanian (TSA)

TSA (UD Treebank for Standard Albanian) është korpusi i parë për shqipen standarde i përfshirë në kornizën UD. Ky korpus është relativisht i vogël, i përbërë nga 60 fjali dhe 922 fjalë në total të mbledhura nga Wikipedia dhe të etiketuara me temën e fjalës, etiketat e pjesëve të ligjëratës, veçoritë morfologjike dhe lidhjet sintaksore (Toska, Nivre, & Zeman, 2020). Përzgjedhja e fjalive është bërë në mënyrë manuale duke synuar përfshirjen e strukturave të larmishme gjuhësore.

Pavarësisht nga përmasat e tij të vogla, ky korpus ka një vlerë të veçantë duke qenë se për herë të parë gjuha shqipe përfshihet në një platformë ndërkombëtare si UD. Në Tabelën 6.2 paraqiten disa statistika kryesore të lidhura me korpusin TSA dhe frekuencën e gjendjes së etiketave të caktuara në këtë bashkësi të dhënash. Në total në këtë korpus janë përdorur 14 etiketa të pjesëve të ligjëratës dhe 33 etiketa sintaksore nga korniza e UD. TSA ka shërbyer si pikënisje e rëndësishme për përfshirjen e shqipes në ekosistemin shumëgjuhësh të UD-së.

Tabela 6.2. Statistika mbi korpusin TSA

Numri i fjalive		60	
Numri i fjalëve		922	
Numri i *shprehjeve shumëfjalëshe* (angl. *multiword tokens*)		-	
Gjatësia mesatare e fjalive		15.37	
Frekuenca e etiketave			
UPOS	**Frekuenca**	**Deprel**	**Frekuenca**
NOUN	238	punct	85
PUNCT	85	det	91
DET	116	case	80
VERB	85	advmod	36
ADP	85	nsubj	68
PRON	53	nmod	29

6.2.2.2. UD Gheg Pear Stories (GPS)

UD Gheg Stories (GPS) është një korpus i cili përmban rirrëfime të videove ‘Pear Stories’ të Wallace Chafe nga folës amtarë të gegërishtes që jetojnë në Zvicër ose nga folës nga Prishtina (Ebert, et al., 2022). Ky korpus përmban 966 fjali, rreth 16000 fjalë nga 64 regjistrimet e marra në konsideratë. Mbledhja e të dhënave është zhvilluar në kuadër të një projekti të zhvilluar në Zyrih, Prishtinë dhe Mynih gjatë periudhës 2019-2022 (Ebert, et al., 2022).

Ky korpus përfaqëson një varietet gjuhësor të rëndësishëm për studimin e shqipes, pasi përfshin një dialekt të gjuhës shqipe (dialektin Gegë). Të dhënat janë marrë në trajtë zanore dhe janë kthyer në trajtë të shkruar duke përdorur mjetin EXMARaLDA, një sistem për kthimin e korpuseve të folura në trajtë të shkruar (Scmidt & Kai, 2014). Ky korpus ka vlerë të veçantë sepse zgjeron përfaqësimin e shqipes në UD përtej standardit, duke përfshirë edhe të dhëna dialektore të folura.

Tekstet janë etiketuar duke përdorur një model të trajnuar me modulin spaCy (Honnibal, Montani, Landeghem, & Boyd, 2020), në Python. Ky model është trajnuar duke u bazuar në korpusin e krijuar nga Kote et al. (2019) dhe një korpus tjetër me të dhëna për gjuhën shqipe standarde. Të gjitha fjalitë e etiketuara automatikisht janë kontrolluar manualisht për gabime të mundshme.

Tabela 6.3. Statistika mbi korpusin GPS

Numri i fjalive		966	
Numri i fjalëve		15,990	
Numri i *shprehjeve shumëfjalëshe* (angl. *multiword tokens*)		-	
Gjatësia mesatare e fjalive		16.55	
Frekuenca e etiketave			
UPOS	**Frekuenca**	**Deprel**	**Frekuenca**
NOUN	2,473	punct	966
PUNCT	966	det	1,280
DET	750	case	978
VERB	2,680	advmod	1,268
ADP	987	nsubj	1,030
PRON	2,898	nmod	389

Në Tabelën 6.3 paraqiten disa statistika kryesore të lidhura me korpusin GPS dhe frekuencën e gjendjes së etiketave të caktuara në këtë bashkësi të dhënash. Në total në këtë korpus janë përdorur 15 etiketa të pjesëve të ligjëratës dhe 35 etiketa sintaksore nga korniza e UD.

6.2.2.3. Saarbrucken Treebank of Albanian Language (STAF)

STAF (Saarbrucken Treebank of Albanian Language) është pjesë e një korpusi paralel 'Corpus of Indo-European Prose and more (CIEP+)' (Verkerk & Talamo, 2024), ku qëllimi kryesor është përfshirja e 18 teksteve letrare të përkthyera në 50 gjuhë nga 12 familje të ndryshme gjuhësore. Ky korpus paralel është etiketuar automatikisht duke përdorur librarinë e Standford Stanza (Qi, Zhang, Zhang, Bolton, & Manning, 2020) të trajnuar duke përdorur korpusin SALT (Kote, et al., 2024). Fjalitë janë etiketuar automatikisht nga modeli i STANZA dhe më pas janë kontrolluar manualisht nga tre folës amtarë të gjuhës shqipe.

STAF përbëhet nga 3325 fjalë, 200 fjali dhe përfshin etiketimet e lidhjeve sintaksore, etiketat e pjesëve të ligjëratës, veçoritë morfologjike dhe temën. Ky korpus është ndërtuar duke u bazuar në rregullat dhe formatin e kornizës së Universal Dependencies (UD). Fjalitë e përfshira në këtë korpus janë marrë nga nëntë libra letrarë të publikuar midis 1963 dhe 2004. Në Tabelën 6.4 paraqiten disa statistika kryesore të lidhura me korpusin STAF dhe frekuencën e gjendjes së etiketave të caktuara në këtë bashkësi të dhënash. Në total në këtë korpus janë përdorur 17 etiketa të pjesëve të ligjëratës dhe 38 etiketa sintaksore nga korniza e UD. STAF ka interes të veçantë sepse mbështetet në tekste letrare dhe në një mjedis paralel shumëgjuhësh, gjë që e bën të dobishëm edhe për studime krahasuese.

Tabela 6.4. Statistika mbi korpusin STAF

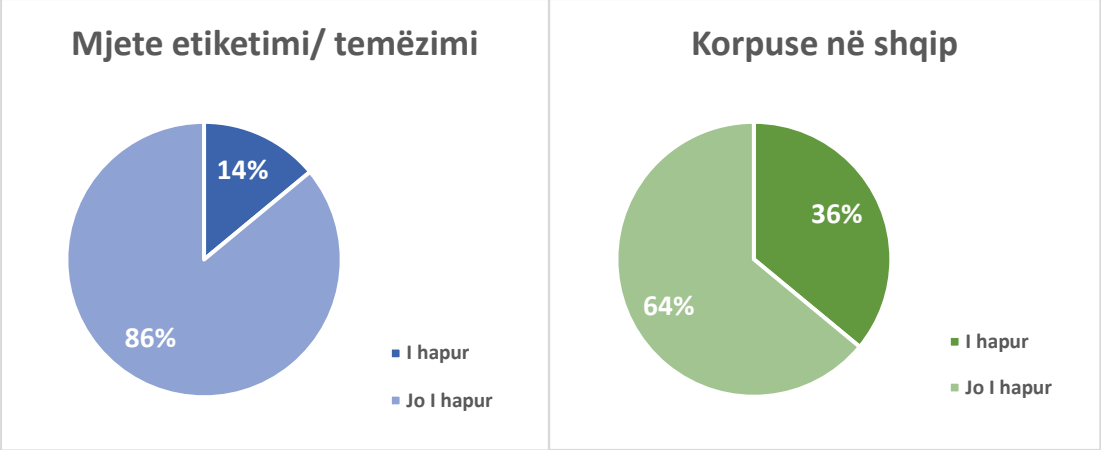
Numri i fjalive		200	
Numri i fjalëve		3325	
Numri i *shprehjeve shumëfjalëshe* (angl. *multiword tokens*)		62	
Gjatësia mesatare e fjalive		17.82	
Frekuenca e etiketave			
UPOS	**Frekuenca**	**Deprel**	**Frekuenca**
NOUN	625	punct	492
PUNCT	492	det	223
DET	299	case	270
VERB	430	advmod	219
ADP	250	nsubj	209
PRON	431	nmod	65

6.2.3. Vlera e sinjaleve konjitive për përpunimin e gjuhëve me burime të kufizuara: rasti i gjuhës shqipe

Pavarësisht nga ekzistenca e disa korpuseve për gjuhën shqipe, ajo mbetet një gjuhë me burime të kufizuara në fushën e përpunimit të gjuhës natyrore, pasi numri i korpuseve është i limituar, si edhe korpuset ekzistuese janë me përmasa jo të mëdha. Mungesa e burimeve paraqet sfida për zhvillimin e modeleve për etiketimin automatik të pjesëve të ligjëratës, analizën sintaksore, semantike dhe nivele të tjera të përpunimit gjuhësor. Ndërkohë, zgjerimi i korpuseve ekzistuese përfshin një angazhim të madh nga ekspertët gjuhësorë, si në kohë ashtu edhe në bashkëpunim të vazhdueshëm ndërmjet specialistëve.

Pjesa më e madhe e modeleve ekzistuese për etiketimin e pjesëve të ligjëratës apo etiketuesve morfologjikë për shqipen nuk janë të disponueshëm online për përdorim apo përmirësime të mëtejshme (Kastrati & Biba, 2022). Ndërkohë, në Grafikun 6.1 paraqitet disponueshmëria e mjeteve dhe korpuseve në gjuhën shqipe, sipas rishikimit të literaturës nga Kote et al. (2022).

Në këtë kontekst, përdorimi i sinjaleve konjitive të marra nga eksperimentet e gjurmimit të syrit apo të mausit përbën një burim alternativ të dhënash të cila ose mund të kompensojnë mungesat e të dhënave gjuhësore ose mund të zgjerojnë informacionin aktual. Duke qenë se të dhënat konjitive përmbajnë jo vetëm informacion gjuhësor, por pasqyrojnë mënyrën si një lexues përpunon gjuhën në kohë reale, mund të konsiderohen një burim tejet i vlefshëm për përdorim në rastin e gjuhëve me pak burime. Këto të dhëna nuk e zëvendësojnë etiketimin gjuhësor manual, por e plotësojnë atë me informacion mbi përpunimin real të tekstit nga lexuesi.



Grafiku 6.1. Niveli i disponueshmërisë së mjeteve të PGJN-së dhe korpuseve në gjuhën shqipe

Një nga arsyet kryesore të përdorimit të të dhënave konjitive informacioni shtesë i ofruar në modelet ekzistuese për të mësuar ndarje më të imëta mes kategorive gramatikore (si p.sh. në rastet kur kategori të caktuara janë të papërcaktuara mirë ose jo mirë të përfaqësuara në të dhënat e korpusit). Së dyti, të dhënat konjitive mund të ndihmojnë në rritjen e aftësisë përgjithësuese të modeleve, duke qenë se korpuset mund të jenë specifike dhe të lidhura me një fushë të caktuar. Kjo është veçanërisht e rëndësishme për shqipen, ku korpuset shpesh janë të kufizuara si nga madhësia, ashtu edhe nga larmia e zhanreve.

6.2.4. Përdorimi i modeleve të paketës Stanza për etiketimin automatik të teksteve në shqip

Stanza është një paketë (angl. *package*) në Python e përdorur për analizën e gjuhëve. Ajo përmban një bashkësi mjetesh të cilat mund të përdoren për analiza të ndryshme gjuhësore. P.sh. mund të përdoret si një pipeline për të konvertuar tekstet në një listë fjalish dhe fjalësh, për të gjeneruar temën e fjalëve, etiketat e pjesëve të ligjëratës, veçoritë morfologjike dhe veçoritë sintaksore për më shumë se 70 gjuhë (Qi, Zhang, Zhang, Bolton, & Manning, 2020). Stanza është ndërtuar përmes komponentëve shumë të saktë të rrjeteve nervore të cilët mundësojnë një trajnim dhe vlerësim shumë të mirë edhe në rastin kur ne duan të punojmë me të dhënat tona. Të gjitha modulet janë

ndërtuar mbi librarinë PyTorch (Paszke, et al., 2019). Në Figura 6.2 paraqitet një përmbledhje e të gjithë pipeline-t të rrjetit neural të përdorur nga Stanza.

Deri tani, gjuha shqipe nuk është pjesë zyrtare e modeleve të disponueshme në paketën e Stanza. Për qëllimet e këtij punimi është përdorur korpusi SALT (Kote, et al., 2024) për trajnimin e modelit, duke ndjekur hapat zyrtare të përcaktuara në faqen zyrtare të librarisë Stanza.

Ky model i trajnuar me Stanza është përdorur për të etiketuar në mënyrë automatike të gjitha tekstet e përfshira në eksperimentet e gjurmimit të syrit apo të mausit. Pas trajnimit të modelit është krijuar pipeline specifik për gjuhën shqipe për të gjitha nivelet e etiketimit:

```
stanza.Pipeline("sq", dir="C:/Users/PC/sq_resources",

download_method=None, allow_unknown_language=True)
```

Duke përdorur funksionin `CoNLL.write_doc2conll()` të përfshirë në `stanza.utils.conllu` janë konvertuar të gjithë filet nga `.txt` në formatin `doc` të Stanza dhe më pas në formatin `CoNLL-u`.

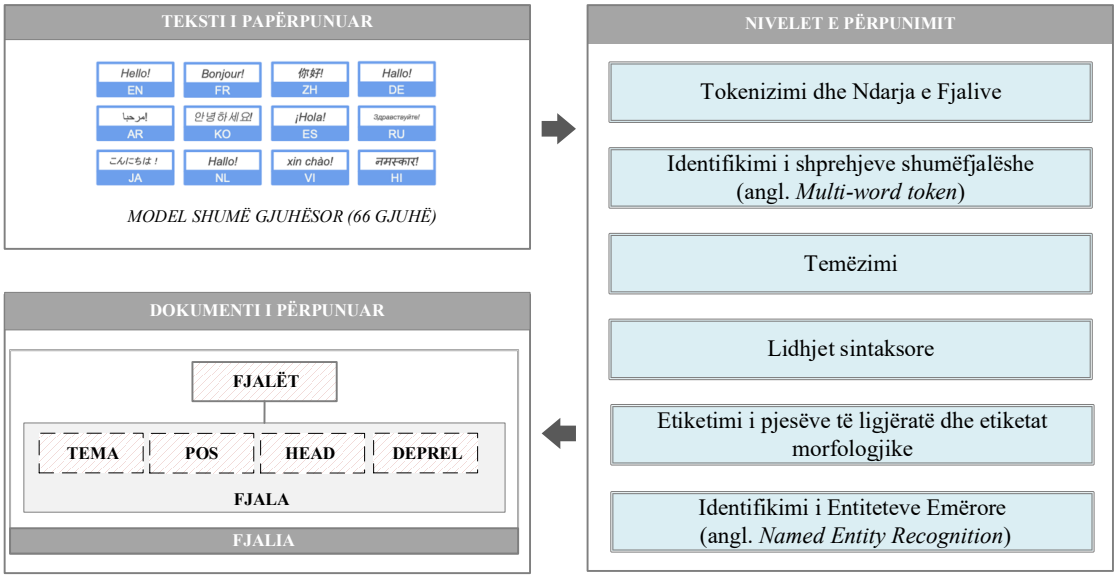


Figura 6.2. Arkitektura e pipeline-it të PGJN-së për rrjetin nervor në paketën Stanza

Në vijim, skedarët CoNLL-u të gjeneruar iu nënshtruan një procesi të detajuar verifikimi nga dy eksperte të gjuhësisë, me qëllim identifikimin e problematikave të mundshme në çdo nivel etiketimi. Të gjithë problematikat e evidentuara u korrigjuan manualisht në platformën ArboratorGrew. Ky proces siguroi një nivel të lartë saktësie në të dhënat e anotuara, duke rritur besueshmërinë e tyre për përdorimin në eksperimente të mëtejshme.

Formati CoNLL-u përmban etiketat e mëposhtme:

- ID: Indeksi i fjalës në fjali.
- FORM: Forma e fjalës, ose simbolit në rastin e shenjave të pikësimit.
- LEMMA: Lema e formës së fjalës.
- UPOS: Etiketa universale e pjesës së ligjëratës.
- FEATS: Veçoritë morfologjike të ndërlidhura me rregullat e gjuhës shqipe.
- HEAD: Lidhja me *rrënjën* (angl. *root*) e fjalisë (zero për rrënjën).
- DEPREL: Lidhja sintaksore në varësi të etiketës HEAD.

Ndërkohë lista e etiketave PoS dhe etiketat përkatëse të lidhura me veçoritë morfologjike janë të paraqitura në Tabelën 6.5 (Kote, et al., 2024).

Tabela 6.5. Lista e etiketave të pjesëve të ligjëratës dhe veçorive morfologjike në korpusin SALT

PJESA E LIGJËRATËS	**ETIKETA PoS**	**VEÇORITË MORFOLOGJIKE**
verb	VERB/AUX	mood, time, person, number, voice; *verb form only in case of participle
noun	NOUN	gender, number, case, definiteness
proper noun	PROPN	gender, number, case, definiteness; Abbr in case of abbreviation
adjective	ADJ	gender, number, case, degree
pronoun	PRON	depends on the type (case, number, gender, person, prontype)
adverb	ADV	AdvType
numeral	NUM	NumType
interjection	INTJ	—
preposition	ADP	case
particle	PART	—
conjunction	CCONJ/SCONJ	—
articles	DET	gender, number, case and prontype
symbols	SYM	—
punctuation marks	PUNCT	—
others	X	Abbr in case of abbreviation

Në Figurën 6.3 paraqiten disa shembuj fjalish nga bashkësitë e të dhënave të përfituara nga eksperimentet e gjurmimit të mausit dhe të syrit, të cilat janë të etiketuara automatikisht duke përdorur modelin Stanza për gjuhën shqipe. Nëpërmjet këtyre shembujve ilustrohet procesi i segmentimit në fjali dhe njësi leksikore, si edhe përcaktimi i kategorive gramatikore dhe ndërtimi i strukturave të varësive sintaksore.

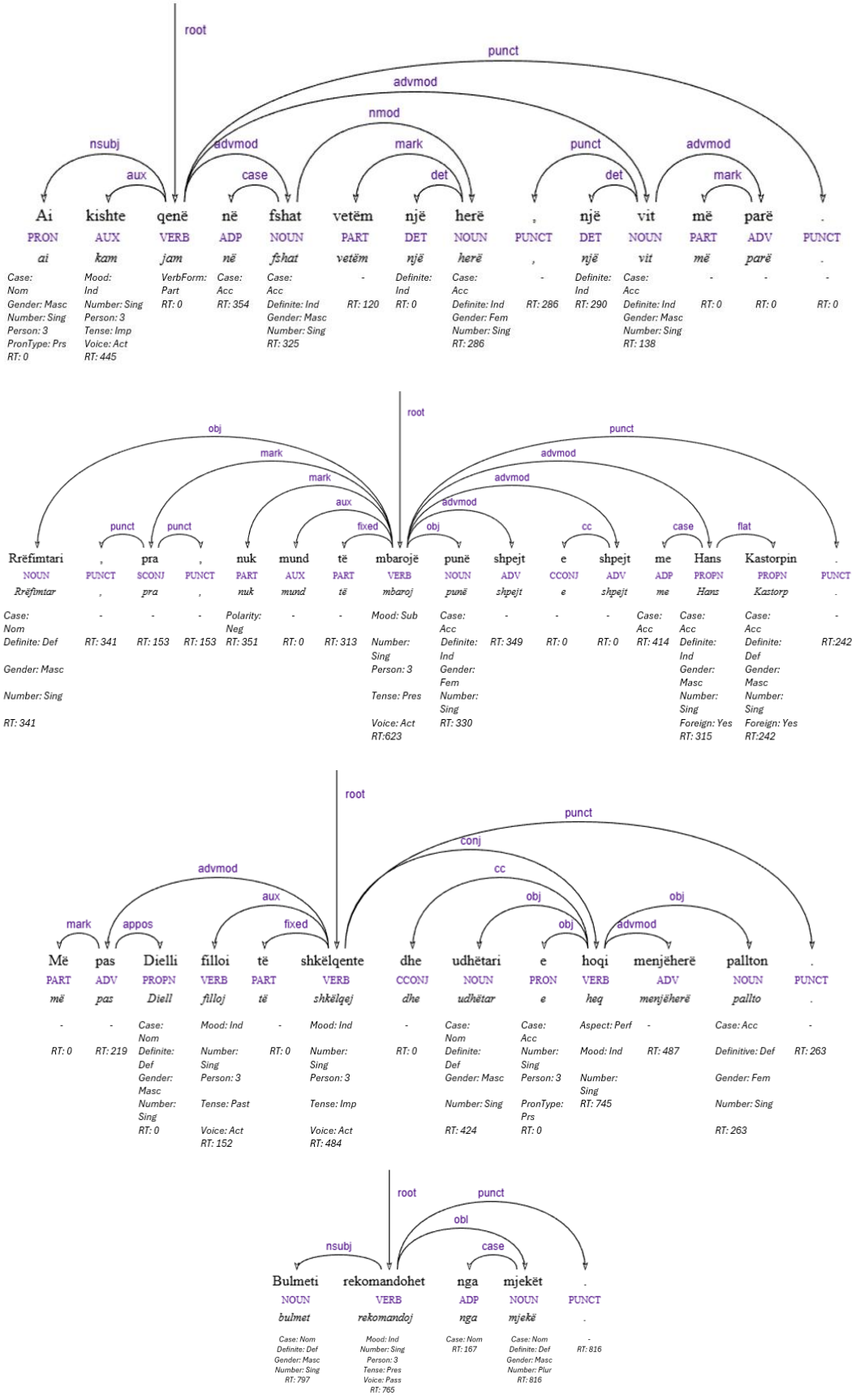


Figura 6.3. Shembuj të etiketimit morfosintaksor të fjalive në eksperimentet e gjurmimit të syrit dhe të mausit

6.3. Integrimi i veçorive gjuhësore dhe të dhënave konjitive për parashikimin automatik të pjesëve të ligjëratës dhe lidhjeve sintaksore

Duke konsideruar kohën dhe burimet e mëdha që kërkohen për të etiketuar manualisht korpuse të mëdha me të dhëna, të dhënat gjatë leximit janë provuar të jenë një alternativë e mirë për të përshpejtuar procesin e etiketimit. Duke marrë në konsideratë përmirësimet e modeleve të PGJN-së nga shtimi i të dhënave për gjurmimin e syrit, të paraqitura në kapitullin 3.2, në këtë pjesë do të paraqesim studimin e përdorimit të të dhënave nga gjurmimi i mausit dhe i syrit për të parashikuar pjesët e ligjëratës dhe lidhjet sintaksore në tekstet në gjuhën shqipe.

Rezultatet e arritura tregojnë se shtimi i të dhënave të leximit nga eksperimentet me gjurmimin e kursorit të mausit apo syrit përmirësojnë dukshëm performancën e modeleve të ndryshme të parashikimit. Këto rezultate theksojnë se bashkimi i të dhënave nga gjurmimi i mausit apo i syrit me tekstet e etiketuara manualisht është një zgjidhje alternative për të përmirësuar performancën e modeleve për gjuhët me burime të pakta si gjuha shqipe. Integrimi i këtyre sinjaleve i afron modelet me mënyrën reale të përpunimit njerëzor të tekstit.

6.3.1. Integrimi i veçorive gjuhësore me bashkësinë e të dhënave nga gjurmimi i mausit dhe i syrit

Para se të mund të bashkojmë të dhënat e mbledhura nga eksperimentet e gjurmimit të mausit dhe syrit duhet të sigurohemi të trajtojmë çdo mospërputhje që mund të ekzistojë midis këtyre bashkësive të të dhënave. Si fillim, duke qenë se të dhënat nga eksperimentet konjitive janë në formatin `.csv`, edhe të dhënat gjuhësore u shndërruan nga formati gjuhësor `.conllu` në formatin `.csv`. Në këtë pikë kemi skedarët e organizuar si më poshtë:

```
Reading_metrics_EyeTracking
-----reading_metrics_001.csv
....
-----reading_metrics_025.csv

Reading_metrics_MouseTracking
-----reading_metrics_001.csv
...
-----reading_metrics_051.csv

CoNLLU_Files
-----PopSci_MultiplEYE_1.conllu
-----PopSci_Caveman_12.conllu
```

```
-----Ins_LearningMobility_3.conllu
-----Ins_HumanRights_2.conllu
-----Lit_MagicMountain_6.conllu
-----Lit_BrokenApril_9.conllu
-----Lit_Alchemist_4.conllu
-----Lit_NorthWind_7.conllu
-----Lit_Solaris_8.conllu
-----Arg_PISACowsMilk_10.conllu
-----Arg_PISARapaNui_11.conllu
-----Enc_WikiMoon_13.conllu

Linguistic_Annotations
-----multiword_tokens.csv
-----all_texts.csv
```

Organizimi i skedarëve në formën e mësipërm reflekton një ndarje të qartë sipas burimit dhe rolit të të dhënave në procesin eksperimental. Në skedarët 'Reading_metrics_EyeTracking' dhe 'Reading_metrics_MouseTracking' ruhen përkatësisht metrikat e leximit të mbledhura për secilin pjesëmarrës. Ndërkohë, në skedarët 'CoNLLU_Files' përfshihen tekstet e etiketuara në nivel gjuhësor, të cilat shërbejnë si bazë për analizat sintaksore dhe semantike, ndërsa në 'Linguistic_Annotations' qëndrojnë burime shtesë për standardizimin e etiketimeve, si lista e shprehjeve shumëfjalëshe dhe lista e plotë e të gjithë fjalëve të etiketuara.

Një vëmendje të veçantë i kemi kushtuar formave klitike të shqipes, si *t'u, t'i, t'ju*, të cilat paraqesin bashkime të pjesëzave funksionale me trajtat e shkurtra të përemrave vetorë ose bashkime vetëm të trajtave të shkurtra. Sipas standardit të platformës UD këto forma analizohen si token shumëfjalësh (angl. *multiword tokens*) dhe ndahen në dy pjesë të veçanta në nivel morfosintaksor, si në shembullin e paraqitur në Figurën 6.4, ku tokeni 14 (*të*) bashkohet me token-in 15 (*e*) në trajtën 14-15 (*ta*).

Në bashkësinë e të dhënave nga eksperimentet e gjurmimit të mausit dhe të syrit, këto forma perceptohen si një fjalë e vetme dhe shoqërohen me një vlerë të vetme kohore leximi. Duke qenë se kjo shkakton një mospërputhje mes njësive konjitive dhe atyre gjuhësore lindi nevoja e identifikimit dhe trajtimit të këtyre rasteve me qëllim integrimin e saktë të të dhënave nga vështrimi me të dhënat gjuhësore nga formati CoNLL-u. Një tjetër problematikë e identifikuar ishte përdorimi i pjesëzës mohuese **s'** si formë e shkurtuar e pjesëzës **nuk**. Edhe në këtë rast, ngjashëm me format klitike, u vendos ndarja në dy pjesë, ku koha e leximit përcaktohet si e njëjtë për të dyja pjesët. Kjo ndarje është bërë për të ruajtur përputhshmërinë metodologjike me standardin e etiketimit të Universal Dependencies (UD), i cili i trajton këto forma si njësi shumëfjalëshe në nivelin morfosintaksor.

Gjatë fazës së parapërpunimit të bashkësive të të dhënave, para bashkimit me të dhënat gjuhësore, u identifikuan të gjitha trajtat e tilla në secilin tekst, të paraqitura në Tabelën 6.6.

Tabela 6.6. Identifikimi i shprehjeve shumëfjalëshe (angl. *multiword tokens*) në tekstet eksperimentale

Forma fillestare	**Fjala 1**	**Fjala 2**	**Etiketat PoS**
iu	i	u	PRON+PRON
t'u	të	u	PART/PRON+PRON
t'i	të	i	PART/PRON+PRON
ta	të	e	PART/PRON+PRON
ia	i	e / i	PRON+PRON
t'ju	të	ju	PART/PRON+PRON
s'ke	s'	ke	PART+AUX
s'mund	s'	mund	PART+AUX
s'duket	s'	duket	PART+AUX
s'kishin	s'	kishin	PART+AUX

Një pikë tjetër e rëndësishme që është konsideruar është përfshirja e shenjave të pikëzimit. Në rastet e eksperimenteve të gjurmimit të mausit apo syrit, tokenat ndahen sipas hapësirave boshe, duke sjellë bashkimin e shenjave të pikëzimit me fjalën përkatëse. Duke qene se marrja në konsideratë e përdorimit dhe pozicionimit të shenjave të pikësimit është tepër e rëndësishme gjatë studimeve gjuhësore, shenjat e pikësimit që ndodheshin të ngjitura në fillim apo fund të një fjale janë ndarë automatikisht si një token më vete, duke marrë të njëjtat metrika leximi me fjalën përkatëse.

Kujdes i veçantë është bërë në disa raste, si p.sh. trajtimin e shkurtesave apo emërtimeve të ndryshme. Në rastin e teksteve tona janë identifikuar disa terma ku nuk ishte e nevojshme ndarja e shenjave të pikësimit, pasi edhe në etiketimin gjuhësor ato trajtohen si njësi e vetme leksikore. Rasti të tilla përmendim termat: 'Erasmus+', 'Inc.', 'Dr.', 'A.G.', 'doc.', '100 000', të cilat nuk janë ndarë gjatë fazës së përpunimit.

```
# text = Shkrimtari Besian Vorpsi me gruan e tij të re, Dianën, vendosën ta bëjnë muajin e mjaltit në Rrafshin e Veriut!
1       Shkrimtari      shkrimtar       NOUN    _       Case=Nom|Definite=Def|Gender=Masc|Number=Sing   13      nsubj   _       _
2       Besian  Besian  PROPN   _       Case=Nom|Definite=Ind|Gender=Masc|Number=Sing   1       appos   _       _
3       Vorpsi  Vorpsi  PROPN   _       Case=Nom|Definite=Def|Gender=Masc|Number=Sing   2       flat    _       _
4       me      me      ADP     _       Case=Acc        5       case    _       _
5       gruan   grua    NOUN    _       Case=Acc|Definite=Def|Gender=Fem|Number=Sing    1       nmod    _       _
6       e       e       DET     _       Case=Acc|Gender=Fem|Number=Sing|PronType=Art    7       det:poss        _       _
7       tij     tij     PRON    _       Case=Acc|Gender=Masc|Number=Sing|Person=3|Poss=Yes|PronType=Prs 5       amod:poss       _       _
8       të      të      DET     _       Case=Acc|Gender=Fem|Number=Sing|PronType=Art    9       det     _       _
9       re      re      ADJ     _       Case=Acc|Degree=Pos|Gender=Fem|Number=Sing      5       amod    _       SpaceAfter=No
10      ,       ,       PUNCT   _       _       11      punct   _       _
11      Dianën  Diana   PROPN   _       Case=Acc|Definite=Def|Gender=Fem|Number=Sing    5       appos   _       SpaceAfter=No
12      ,       ,       PUNCT   _       _       11      punct   _       _
13      vendosën        vendos  VERB    _       Mood=Ind|Number=Plur|Person=3|Tense=Past|Voice=Act      0       root    _       _
14-15   ta      _       _       _       _       _       _       _
14      të      të      PART    _       _       16      mark    _       _
15      e       e       PRON    _       Case=Acc|Gender=Masc|Number=Sing|Person=3|PronType=Prs  16      obj     _       _
16      bëjnë   bëj     VERB    _       Mood=Ind|Number=Plur|Person=3|Tense=Pres|Voice=Act      13      xcomp   _       _
17      muajin  muaj    NOUN    _       Case=Acc|Definite=Def|Gender=Masc|Number=Sing   16      obj     _       _
18      e       e       DET     _       Case=Acc|Gender=Masc|Number=Sing|PronType=Art   19      det:poss        _       _
19      mjaltit mjaltë  NOUN    _       Case=Gen|Definite=Def|Gender=Masc|Number=Sing   17      nmod    _       _
20      në      në      ADP     _       Case=Acc        21      case    _       _
21      Rrafshin        rrafsh  NOUN    _       Case=Acc|Definite=Def|Gender=Masc|Number=Sing   16      obl     _       _
22      e       e       DET     _       Case=Acc|Gender=Masc|Number=Sing|PronType=Art   23      det:poss        _       _
23      Veriut  veri    NOUN    _       Case=Gen|Definite=Def|Gender=Masc|Number=Sing   21      flat    _       SpaceAfter=No
24      !       !       PUNCT   _       _       13      punct   _       SpaceAfter=No
```

Figura 6.4. Shembull i formave klitike në tekstet e përfshira në eksperimentet e gjurmimit të syrit dhe të mausit

Me rregullimin e të gjitha trajtave problematike është krijuar një skript kontrolli i cili kontrollon përputhjen e fjalëve një nga një, përkatësisht në dy bashkësitë e të dhënave: të dhënat konjitive nga gjurmimi i mausit/syrit dhe të dhënat nga etiketimi gjuhësor. Në të dyja bashkësitë u krijua një kolonë e re `word_nr_text`, e cila mban pozicionimin e secilës fjalë në secilin tekst. Të dhënat gjuhësore dhe ato konjitive janë bashkuar në nivel fjale duke përdorur përputhshmërinë mes tre kolonave `text_id`, `word_nr_text` dhe `word`. Kemi të krijuar dy dosje të reja ku ndodhen skedarët e bashkuara (veçori konjitive+ veçori linguistike) për secilin pjesëmarrës në eksperiment:

```
Merged_EyeTracking
-----conllu_gaze_001.csv
....
-----conllu_gaze_025.csv


Merged_MouseTracking
-----conllu_mouse_001.csv
...
-----conllu_mouse_051.csv
```

Pas bashkimit të të dhënave për secilin pjesëmarrës në eksperiment, është krijuar edhe një bashkësi të dhënash ku gjeneralizohen sjelljet e veçanta të secilit pjesëmarrës. Për këtë arsye janë përllogaritur mesatarja dhe devijimi standard për secilën nga metrikat e bashkësisë fillestare të të dhënave nga leximit.

Përtej etiketave gjuhësore, në këtë studim janë marrë në konsideratë edhe veçori të ndërlidhura me strukturën e fjalisë dhe karakteristikat e secilës fjalë. Këto veçori janë përllogaritur në mënyrë automatike për secilën fjalë në bashkësinë e të dhënave në mënyrë që gjatë procesit të trajnimit të konsiderohej edhe informacioni i lidhur me strukturën e tekstit. Në Tabelën 6.7 paraqiten të gjithë këto veçori sipas grupimit përkatës.

Në rastin e bashkësisë së të dhënave nga gjurmimi i mausit, numri total i veçorive është 41 veçori (angl. *features*) për 914 fjalë të përfshira në tekstet eksperimentale. Ndërsa, në rastin e bashkësisë së të dhënave nga gjurmimi i syrit, numri i veçorive është 120 për 7580 fjalë të përfshira në tekstet eksperimentale. Ky dallim reflekton kompleksitetin më të lartë të të dhënave të gjurmimit të syrit krahasuar me ato të mausit. Për më tepër, numri më i madh i veçorive dhe i instancave në të dhënat nga gjurmimi i syrit ofron një potencial më të lartë për ndërtimin e modeleve më të sakta dhe analizave më të thelluara.

Tabela 6.7. Veçoritë strukturore të përfshira në bashkësinë e të dhënave nga gjurmimi i mausit dhe i syrit

Kategoria	***Veçoria* (angl. *feature*)**	**Përshkrimi**	**Tipi i veçorisë**
Veçori të ndërlidhura me strukturën e fjalisë	text_id	Numri identifikues i tekstit	Numerike
	wordId	Pozicionimi i fjalës specifike në tekst	Numerike
	wordId_sentence	Pozicionimi i fjalës në një fjali	Numerike
	sentence_id	Numri i fjalisë në tekstin specifik	Numerike
	is_first	Nëse fjala është apo jo e para në fjali	Boolean
	is_last	Nëse fjala është apo jo e fundit në fjali	Boolean
Veçori ortografike	is_capitalized	Nëse fjala fillon me shkronjë të madhe apo jo	Boolean
	is_all_caps	Nëse fjala është e shkruar e gjitha me shkronja të mëdha	Boolean
	is_all_lower	Nëse fjala është e shkruar e gjitha me shkronja të vogla	Boolean
	is_alphanumeric	Nëse fjala është një kombinim i numrave dhe shkronjave	Boolean
	has_hyphen	Nëse fjala përmban shenjë '-'	Boolean
	is_numeric	Nëse fjala përfaqëson një numër	Boolean
	capitals_inside	Nëse fjala përmban shkronja të mëdha në brendësi të saj	Boolean
Veçori morfologjike	prefix_n (n=1-4)	N shkronjat e para të fjalës ku n konsiderohet nga 1 në 4	String
	suffix_n (n=1-4)	N shkronjat e fundit të fjalës ku n konsiderohet nga 1 në 4	String
Veçori kontekstuale	prev_word	Fjala paraardhëse	String
	next_word	Fjala pasardhëse	String

6.3.2. Analiza e shpërndarjes së kohës së leximit në funksion të veçorive gjuhësore dhe metrikave të leximit

Rayner et al (2007) përcaktojnë se 35% e kohës së leximit përqendrohet në *fjalët funksionale* (angl. *function words*) të karakterizuara nga etiketat DET, CCONJ, PRON, AUX. Në kontrast, rreth 85% e kohës së vështrimit përqendrohet në fjalë me përmbajtje leksikore, të karakterizuara nga etiketat NOUN, VERB, ADJ, ADV. Në Figurën 6.5 paraqitet shpërndarja e kohëzgjatjes së vështrimit (angl. *gaze duration)* sipas 14 etiketave të ndryshme të pjesëve të ligjëratës, në secilin nga tekstet e përfshira në eksperimentet e gjurmimit të mausit. Rezultatet tregojnë se, megjithëse ekzistojnë dallime në shpërndarjen absolute të kohëzgjatjes së leximit midis teksteve të ndryshme, këto ndryshime lidhen kryesisht me veçori stilistike dhe diskursive të tipologjive të ndryshme tekstuale.

Pavarësisht nga ndryshimet, vërehet një prirje ndërmjet teksteve, përqendrimi më i lartë qëndron në fjalët e përmbajtjes. Informacioni i paraqitur në trajtë vizuale në Figurën 6.5 sugjeron se lexuesit i kushtojnë më shumë vëmendje njësive që mbartin informacion semantik dhe sintaksor më kompleks, në përputhje me gjetjet e Rayner et al. (2007).

Figua 6.5, në këtë mënyrë konfirmon faktin se modelet e vëmendjes vizuale gjatë leximit ndjekin parime universale, pavarësisht nga gjuha e përdorur. Kjo vërteton se gjuha shqipe është e krahasueshme në këtë kontekst me gjuhë të tjera më të përdorura, duke motivuar punime të ardhshme ku informacioni konjitiv nga gjuhë të tjera mund të përdoret për të përmirësuar modelet e përpunimit të gjuhës shqipe.

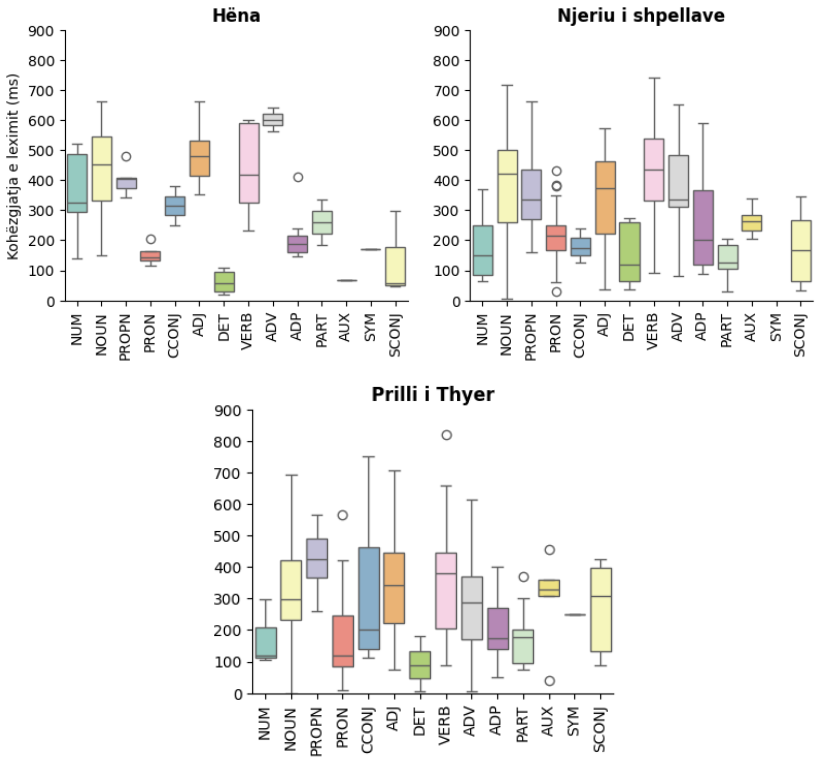


Figura 6.5. Shpërndarja e kohës së leximit për secilën etiketë PoS në eksperimentet e gjurmimit të kursorit të mausit

Në Figurën 6.6 paraqitet shpërndarja e kohëzgjatjes së leximit (në milisekonda) për secilën kategori gramatikore (UPOS), e ndarë sipas teksteve të përfshira në eksperimentet e gjurmimit të syrit. Boxplotet e paraqitura tregojnë medianën, intervalin interkuantil dhe vlerat ekstreme për secilën etiketë. Rezultatet tregojnë një model të qëndrueshëm në pothuajse të gjitha tekstet, ku kategoritë leksikore të hapura (NOUN, VERB dhe ADJ) shfaqin medianë më të lartë të kohëzgjatjes së leximit krahasuar me kategoritë funksionale (DET, ADP, SCONJ, CCONJ, PART dhe AUX). Ky dallim në kohëzgjatje është në përputhje me literaturën, ku fjalët me ngarkesë më të lartë semantike kërkojnë më shumë burime konjitive për përpunim.

Në të kundërt kategoritë funksionale paraqesin kohë leximi më të ulëta dhe shpërndarje më të ngushta, e cila sugjeron një përpunim më automatik. P.sh., etiketat DET dhe ADP në shumicën e teksteve karakterizohen nga medianë të ulët dhe variancë të kufizuar. Kjo shpjegohet me faktin se këto njësi gjuhësore kanë rol kryesisht strukturor. Megjithatë në figurë vërejmë një variabilitet midis teksteve të ndryshme. Disa tekste paraqesin një shpërndarje më të gjerë të kohëzgjatjes së leximit për të njëjtat kategori gramatikore, duke sugjeruar ndikimin e faktorëve të diskursit dhe stilistikës.

Në kuadër të integrimit të të dhënave nga eksperimentet e gjurmimit të syrit në modelet e përpunimit automatik të gjuhës, këto rezultate theksojnë akoma më shumë se veçoritë konjitive të lidhura me kohëzgjatjen e leximit janë veçanërisht informuese për dallimin mes kategorive leksikore dhe atyre funksionale. Kjo është e lidhur drejtpërdrejt me modelimin e etiketuesve automatik të pjesëve të ligjëratës të cilët do të studiohen në vijim të këtij kapitulli. Figura 6.6 demonstron se shpërndarja e kohës së leximit përmban të dhëna të rëndësishme të cilat mund të përmirësojnë performancën e klasifikimit.

Në vijim, në Figurën 6.7 paraqitet shpërndarja e kohëzgjatjes së leximit (në milisekonda) për secilën tip të marrëdhënieve sintaksore të varësisë (etiketa deprel), e ndarë sipas teksteve të përfshira gjithashtu në eksperimentet e gjurmimit të syrit. Rezultatet tregojnë se elementët e strukturës së fjalisë , si nsubj, obj, ccomp dhe xcomp, shfaqin medianë më të lartë të kohëzgjatjes së leximit krahasuar me marrëdhëniet më periferike ose funksionale. Marrëdhëniet si det, case, cc dhe mark paraqesin kohë më të ulëta leximi dhe shpërndarje më të ngushta. Këto elemente kanë funksion kryesisht gramatikor dhe shërbejnë për të sinjalizuar marrëdhënie strukturore, por nuk mbartin ngarkesë të lartë semantike.

Edhe në këtë rast vërejmë ndryshime mes teksteve të ndryshme. Tekstet të cilat kanë një kompleksitet më të lartë paraqesin shpërndarje më të gjera të kohës së leximit për marrëdhënie të caktuara. Nga perspektiva e trajnimit të modeleve kompjuterike, shpërndarja e paraqitur në Figurën 6.7 tregon se kohëzgjatja e leximit ndryshon sistematikisht sipas llojit të marrëdhënieve sintaksore, duke sjellë që metrikat e nxjerra nga eksperimentet e gjurmimit të syrit të mund të përdoren si veçori (angl. *feature*) për detyra si parashikimi i HEAD dhe DEPREL. Këto rezultate mbështesin hipotezën se të dhënat konjitive përmbajnë informacion të strukturuar mbi strukturën sintaksore të gjuhës dhe mund të kontribuojnë në përmirësimin e modeleve të inteligjencës artificiale për analizën sintaksore, veçanërisht në rastin e gjuhëve me burime të kufizuara si shqipja.

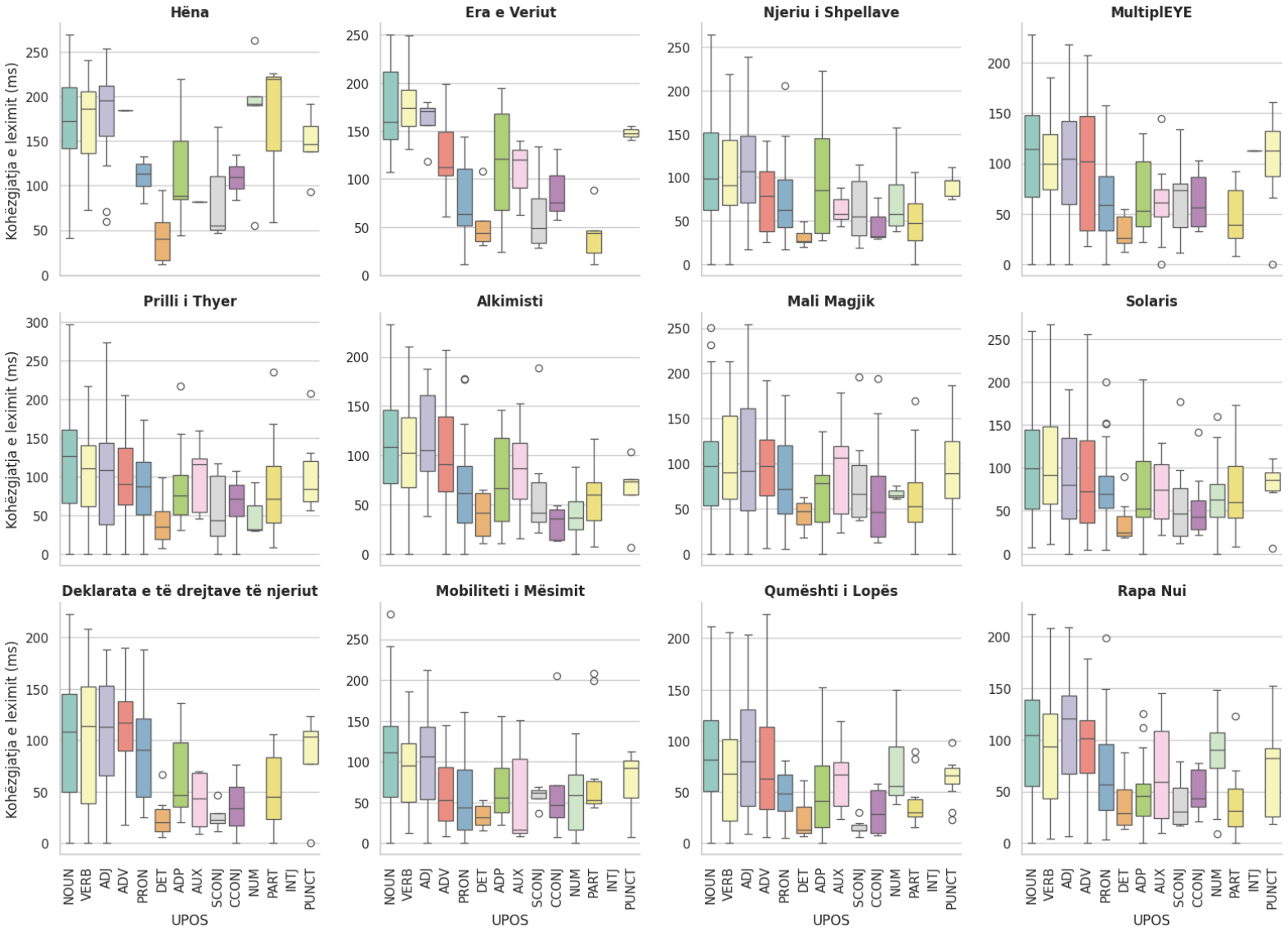


Figura 6.6. Shpërndarja e kohës së leximit për secilën etiketë PoS në eksperimentet e gjurmimit të syrit

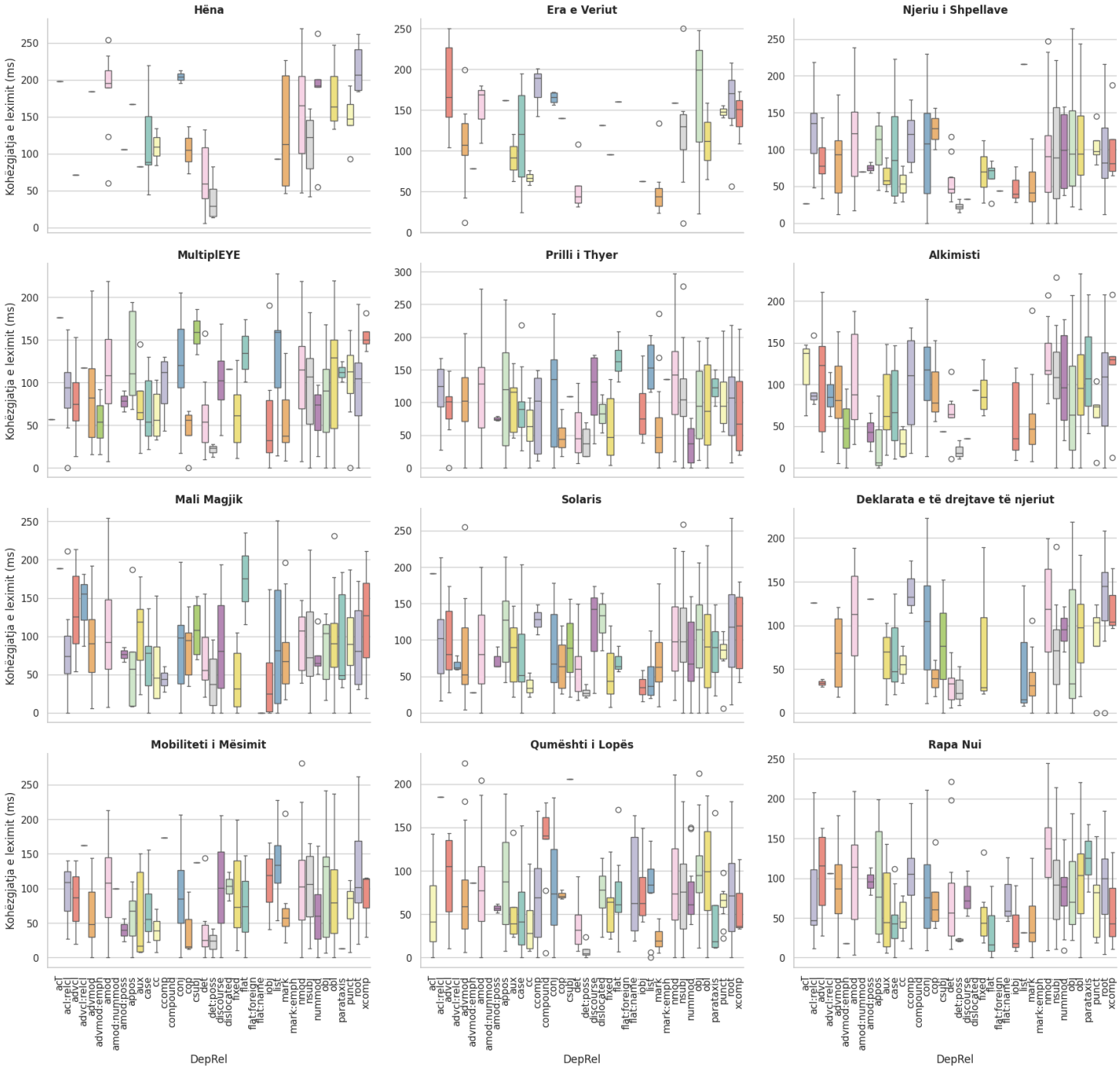


Figura 6.7. Shpërndarja e kohës së leximit për secilën etiketë Deprel në eksperimentet e gjurmimit të syrit

6.3.3. Ndërtimi dhe trajnimi i një modeli CRF mbi veçoritë gjuhësore dhe metrikat e leximit për parashikimin e pjesëve të ligjëratës

Duke u bazuar në veçoritë e përzgjedhura dhe në sasinë e të dhënave tona, modeli që është përzgjedhur për të krijuar etiketuesin e pjesëve të ligjëratës është CRF (Conditional Random Fields) një model statistikor i të nxënit të mbikëqyrur (angl. *supervised machine learning*) i cili mund të konsiderohet një variacion i modelit Markov Random Field. Ky model përllogarit probabilitetin që një bashkësi gjendjesh janë apo jo të varuara nga njëra-tjetra duke u bazuar në një bashkësi observimesh. Në rastin e problemit tonë, modeli do të vlerësojë probabilitetin që një fjalë e observuar në një kontekst të caktuar (ku konteksti përcaktohet nga veçoritë e lartpërmendura) i përket një etikete specifike të pjesëve të ligjëratës. Për vlerësimin e modelit, përdorim metrikën F1-score, si një metrikë që kombinon dy metrikat saktësia (angl. *precision*) dhe ndjeshmëria (angl. *recall*).

Në këtë eksperiment u përdor një kombinim i parametrave si më poshtë:

- c1: parametri i L1 regularization për modelin CRF. Ne kemi eksperimentuar me vlerat [0.001, 0.01, 0.1, 0.5].
- c2: parametri i L2 regularization për modelin CRF. Ne kemi eksperimentuar me vlerat [0.01, 0.1, 0.5, 1.0].
- max_iterations: numri maksimal i iteracioneve për të optimizuar algoritmin. Kemi përdorur vlerën default 100.
- all_possible_transitions: nëse do të përfshihen të gjitha tranzicionet e mundshme në modelin CRF. Vlera default është False, ndërkohë ne kemi përdorur vlerën True.

Për secilin kombinim të c1 dhe c2 u trajnuan dhe testuan 5 modele (një për secilën ndarje të validimit të kryqëzuar). Për secilin nga këto raste u llogarit F1-score me peshë në mënyrë që të merret frekuenca për secilën klasë. Mesatarja e F1-score u ruajt për secilën ndarje në mënyrë që të identifikohen parametrat me performancën më të mirë, të paraqitur në ngjyrë të zezë më të theksuar në Tabelën 6.8.

Nga Tabelën 6.8 rezulton që parametrat më optimal të modelit për bashkësinë e të dhënave nga eksperimentet e gjurmimit të mausit janë $c1 = 0.001$ dhe $c2 = 0.01$, ku modeli arrin një vlerë 0.7687 për metrikën F1-score. Ndërkohë, në rastin e modelit që përdor të dhënat nga gjurmimi i syrit rezultati më i mirë arrihet për parametrat $c1 = 0.001$ dhe $c2 = 0.5$, ku modeli arrin një vlerë 0.9166 për metrikën F1-score.

Ky rezultat justifikohet duke marrë në konsideratë që bashkësia e të dhënave e krijuar nga eksperimentet e gjurmimit të mausit është relativisht e vogël në krahasim me ato të marra nga eksperimentet e gjurmimit të syrit. Megjithatë, nga ky eksperiment arrijmë në përfundim se të dhënat konjitive, edhe në sasi të vogla, arrijnë të diferencojnë mes etiketave të ndryshme të ligjëratës, duke përdorur vetëm informacionin strukturor të fjalisë, temën e fjalës dhe metrikat e leximit. Kjo sugjeron se integrimi i të dhënave nga leximi mund të jetë një burim i vlefshëm informacioni për algoritmet e etiketimit morfologjik dhe sintaksor edhe për gjuhën shqipe.

Tabela 6.8. Rezultatet e F1-score për kombinime të parametrave c1 dhe c2 në modelin CRF për parashikimin e pjesëve të ligjëratës

C1	**C2**	**F1-score Gjurmimi i Mausit**	**F1-Score Gjurmimi i Syrit**
0.001	0.01	**0.7687**	0.9054
0.001	0.1	0.7677	0.9146
0.001	0.5	0.7641	**0.9166**
0.001	1.0	0.7474	0.9145
0.01	0.01	0.7660	0.9068
0.01	0.1	0.7648	0.9146
0.01	0.5	0.7575	0.9150
0.01	1.0	0.7447	0.9145
0.1	0.01	0.7455	0.9081
0.1	0.1	0.7591	0.9142
0.1	0.5	0.7590	0.9154
0.1	1.0	0.7490	0.9132
0.5	0.01	0.7356	0.9109
0.5	0.1	0.7380	0.9123
0.5	0.5	0.7235	**0.9166**
0.5	1.0	0.7078	0.915

Në Figurën 6.8 paraqitet në trajtë vizuale (nëpërmjet një heatmap-i) rezultatet e arritura nga kombinimet e parametrave c1 dhe c2 të algoritmit CRF në rastin e të dy tipeve të eksperimenteve. Modelet CRF me performancën më të lartë u përdorën për të llogaritur F1-score për secilën etiketë të pjesëve të ligjëratës në të dy rastet e eksperimenteve.

Analiza e rezultateve të paraqitura lejon identifikimin më të lehtë të konfigurimeve optimale të parametrave që kontribuojnë në performancën më të mirë të modelit. Për më tepër, ajo ofron një kuptim më të thellë të ndikimit të parametrave në cilësinë e parashikimeve për kategori të ndryshme gramatikore.

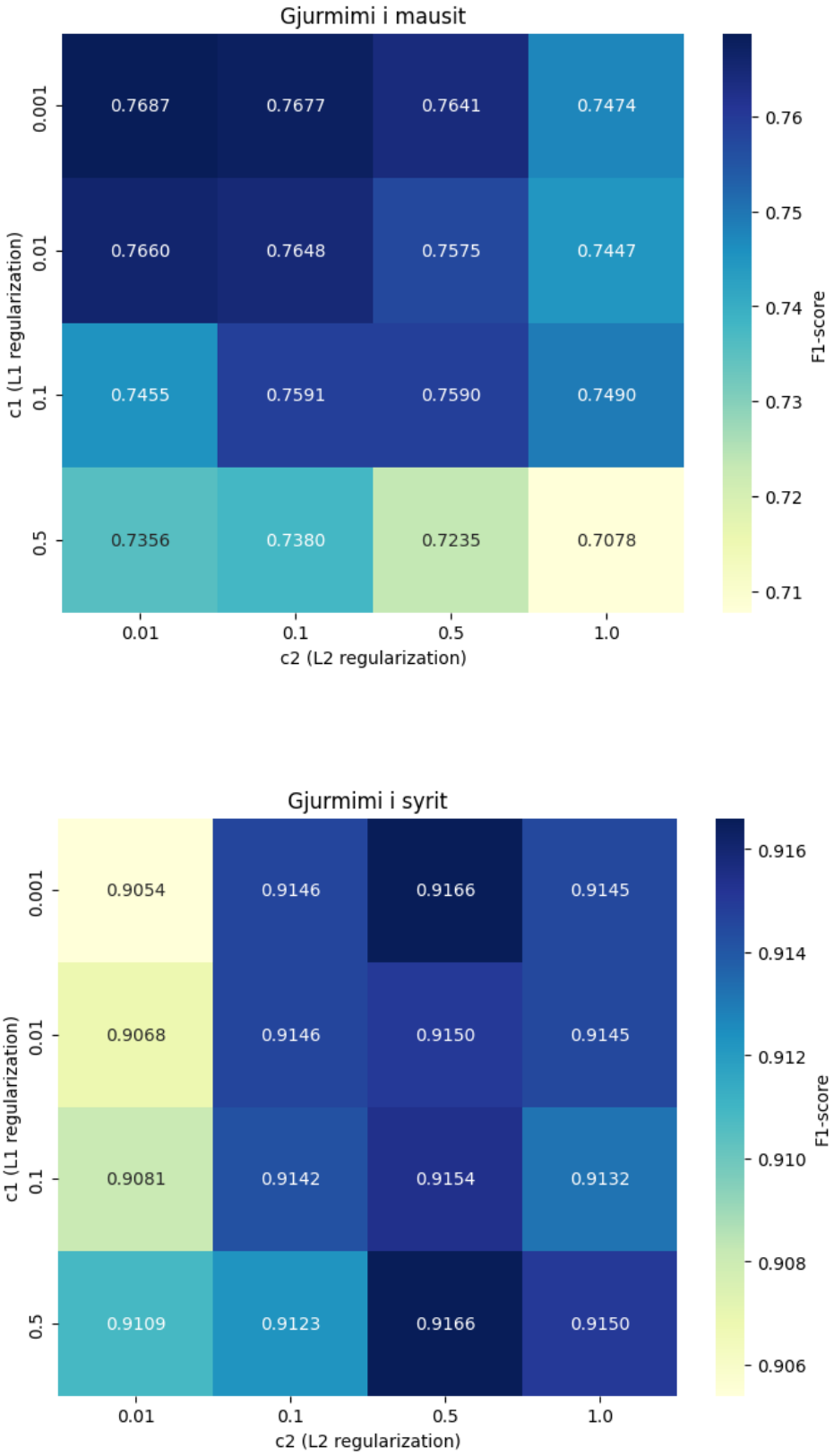


Figura 6.8. Heatmap i F1-score për kombinime të parametrave c1 dhe c2 në modelin CRF për parashikimin e pjesëve të ligjëratës

Në Tabelën 6.9 dhe Tabelën 6.10. paraqiten saktësitë e arritura gjatë parashikimit të secilës etiketë ligjërate, duke bërë ndarjen në tre nivele të ndryshme saktësie. Kjo ndarje mundëson një analizë më të detajuar të performancës së modelit për kategori të ndryshme gramatikore, duke evidentuar etiketat më problematike. Gjithashtu, krahasimi mes dy tipeve të eksperimenteve ofron një pasqyrë të ndikimit të tipit të të dhënave konjitive në rezultatet e parashikimit.

Tabela 6.9. Rezultatet e F1-score për secilën etiketë të pjesës së ligjëratës duke përfshirë metrikat e leximit nga eksperimentet e gjurmimit të mausit

ETIKETA POS	F1-SCORE	NIVELI I SAKTËSISË
ADP	0.9065	Saktësi > 0.80
DET	0.8715	
SCONJ	0.8700	
PROPN	0.8187	
VERB	0.7838	0.70 < Saktësi < 0.80
NOUN	0.7735	
PRON	0.7719	
PART	0.7782	
CCONJ	0.7220	
ADJ	0.6443	Saktësi < 0.7
AUX	0.5921	
NUM	0.5889	
ADV	0.4395	
SYM	0.0000	

Analiza e performancës në nivel etikete gjuhësore tregon se përdorimi i të dhënave nga gjurmimi i syrit sjell përmirësime të dukshme krahasuar me përdorimin e të dhënave nga gjurmimi i mausit. Etiketat më komplekse morfologjikisht, si AUX, ADV dhe ADJ, përfitojnë nga informacioni vizual i leximit duke shënuar përmirësime të larta deri në +0.3362 për AUX dhe +0.2806 për ADV. Korpusi i gjurmimit të syrit është një korpus me kompleks dhe më i zgjeruar se ai i gjurmimit të mausit. Kjo larmishmëri të dhënash lejon modelin të mësojë variacione më të mëdha dhe të identifikojë më saktë klasat e ndryshme gjuhësore.

Tabela 6.10. Rezultatet e F1-score për secilën etiketë të pjesës së ligjëratës duke përfshirë metrikat e leximit nga eksperimentet e gjurmimit të syrit

Etiketa PoS	F1-score	Përmirësimi në krahasim me gjurmimin e mausit	Niveli i saktësisë
PUNCT	0.9959	-	Saktësi > 0.90
DET	0.9835	+0.1120	
ADP	0.9599	+0.0534	
CCONJ	0.9361	+0.2141	
NOUN	0.9304	+0.1569	
AUX	0.9283	+0.3362	
PRON	0.9074	+0.1355	
SCONJ	0.8983	+0.0283	0.80 < Saktësi < 0.90
ADJ	0.8772	+0.2329	
PART	0.8733	+0.0951	
VERB	0.8494	+0.0656	
NUM	0.7853	+0.1964	0.70 < Saktësi < 0.80
PROPN	0.7300	−0.0887	
ADV	0.7201	+0.2806	
SYM	0.0000	0.0000	

Këto rezultate sugjerojnë se informacioni i marrë nga eksperimente konjitive, kur kombinohet me një bashkësi të dhënash të larmishme, ndihmon modelin të dallojë më lehtësisht mënyrën e përdorimit të fjalëve në kontekste të ndryshme, duke rritur saktësinë e klasifikimit për grupe më sfiduese të etiketave të pjesëve të ligjëratës. Përmirësimi nuk lidhet vetëm me shtimin sasior të të dhënave, por edhe me faktin që sinjalet nga gjurmimi i syrit bartin informacion më të imët mbi vëmendjen dhe përpunimin gjuhësor gjatë leximit.

6.3.4. Ndërtimi dhe trajnimi i një modeli të bazuar në arkitekturën BiLSTM për parashikimin e pjesëve të ligjëratës

Duke marrë në konsideratë faktin se të dhënat e mbledhura nga eksperimentet e gjurmimit të syrit janë më të larmishme dhe më të mëdha në numër, si model të dytë për të testuar kemi përzgjedhur një arkitekturë BiLSTM (Bidirectional Long Short-Term Memory). Në këtë rast kemi përdorur bashkësinë e të dhënave të agreguara nga eksperimentet e gjurmimit të syrit, ku përllogariten mesatarja dhe devijimi standard

ndër pjesëmarrës të ndryshëm. Në këtë arkitekturë kemi përfshirë përfaqësimin leksikor të fjalëve (angl. *word embeddings*) dhe veçoritë numerike të nxjerra nga të dhënat e gjurmimit të syrit. Të dhënat janë agreguar në nivel tokeni, duke kombinuar informacionin gjuhësor me metrikat e leximit.

Përveç veçorive bazë, u krijuan edhe disa veçori të reja të normalizuara për të kapur marrëdhëniet strukturore mes lëvizjeve të syrit dhe strukturës së fjalisë. Konkretisht, metrikat e përllogaritura janë proporcioni i numrit të sakadave hyrëse në fjalë me gjatësinë e fjalës (`inword_per_length`), proporcioni i numrit të sakadave dalëse në fjalë me numrin mesatar të fiksimeve në atë fjalë (`outword_per_fix`) dhe raporti i regresioneve (`sakadat hyrëse/(sakada hyrëse+dalëse`). Të gjitha metrikat e leximit janë normalizuar duke aplikuar formulën e *z-score* për secilin tekst, duke përqendruar modelin në variacionet relative brenda tekstit.

Vlerësimi i modelit u bë duke përllogaritur *accuracy*, *macro F1-score* dhe *weighted F1-score* të detajuar për secilën etiketë të pjesëve të ligjëratës. Modeli u testua me pesë konfigurime të ndryshme, të paraqitura në Tabelën 6.11.

Tabela 6.11. Konfigurimet e eksperimenteve për parashikimin e pjesëve të ligjëratës me arkitekturën BiLSTM

KONFIGURIMI	**VEÇORITË**	**SHKURTIMI**
Pa metrika leximi	Lemma+Vecoritë e përshkruara në Tabela 6.7	No Gaze **(NG)**
Me metrikat e leximit bazë	NG+ number_of_fixations, gaze_duration, total_fixation_duration	Base **(B)**
Me metrika bazë + të herëta (early)	NG+B+ first_fixation_duration, first_duration	Basic+Early **(B+E)**
Me metrika bazë + të herëta (early) + të vona (late)	NG+B+E+ re_reading_time, second_pass_duration, go_past_time	Basic+Early+Late **(B+E+L)**
Me metrika bazë + të herëta (early) + të vona (late) + sakadat	NG+B+E+L+ inword_per_length, outgoing_per_fix, regression_ratio	Basic+Early+Late +Saccade **(B+E+L+S)**

Konfigurimi i parametrave të modelit është përcaktuar duke konsideruar ruajtjen e balancës mes kapacitetit përfaqësues të modelit dhe shmangies së mbipërshtatjes (angl. *overfitting*), si edhe duke konsideruar tipin e bashkësisë së të dhënave që kemi në dispozicion. Dimensioni i shtresës embedding u vendos në 100, dimensioni i shtresës së fshehtë (angl. *hidden layer*) u caktua 64, ndërsa dimensioni i veçorive nga gjurmimi

i syrit përcaktohet në varësi të veçorive të përzgjedhura për t'u përdorur gjatë trajnimit të modelit. Procesi i optimizimit u realizua me algoritmin Adam me nivel mësimi (angl. *learning rate)* prej 0.001 dhe batch size me vlerë 32, të cilat janë vlera standarde për arkitekturat neurale të këtij tipi. Gjithashtu, u përdor edhe një funksion humbje *CrossEntropyLoss*. Numri i epokave fillimisht u përzgjodh në 5, duke qenë se modelet BiLSTM në bashkësi të dhënash me përmasa mesatare priren të konvergjojnë shpejt, dhe një numër i lartë epokash do të rriste rrezikun e mbipërshtatjes (angl. *overfitting*).

Tabela 6.12 paraqet performancën e modelit BiLSTM, me vlerat e F1-score, për etiketimin e pjesëve të ligjëratës në pesë konfigurimet eksperimentale të prezantuara në Tabelën 6.11, duke përdorur vlerat mesatare të metrikave të leximit. Rezultatet e paraqitura tregojnë se integrimi i informacionit nga eksperimentet e gjurmimit të syrit ndikon në mënyrë selektive në parashikimin e disa kategorive morfosintaksore.

Tabela 6.12. Rezultatet e F1-score në eksperimentet për parashikimin e pjesëve të ligjëratës me arkitekturën BiLSTM duke përdorur vlerën mesatare

POS Tag	**No Gaze**	**B**	**B+E**	**B+E+L**	**B+E+L+S**
NOUN	0.8592	0.8641	0.8604	**0.8686**	0.8579
PUNCT	0.9383	0.9195	0.9344	0.9327	**0.9412**
AUX	0.8014	**0.8618**	0.8205	0.8388	0.8307
DET	0.8522	0.8577	**0.8659**	0.8610	0.8628
PRON	0.7422	0.7372	0.7928	0.7660	**0.7960**
VERB	0.6067	**0.6603**	0.6323	0.5894	0.6548
CCONJ	0.8331	0.8843	0.8850	**0.9276**	0.8873
ADJ	0.6018	0.6988	0.6890	0.6840	**0.7107**
ADP	0.8775	0.8738	**0.9119**	0.9045	0.8902
PART	0.5734	0.5952	**0.6171**	0.6156	0.6124
ADV	0.6360	**0.7146**	0.6393	0.6652	0.6549
SCONJ	0.8000	0.7674	0.7571	0.7359	**0.8107**
PROPN	0.5705	0.5536	0.5441	**0.5948**	0.5586
NUM	0.5669	0.6069	0.5950	**0.6143**	0.6047

Për shembull, për kategorinë NOUN vërehet një rritje graduale e performancës deri në konfigurimin B+E+L, ndërkohë që vërehet se shtimi i veçorive nga sakadat e lëvizjes së syrit nuk sjell përmirësim të mëtejshëm. Një përmirësim i ndjeshëm vihet re në etiketën VERB, ku modeli arrin të përmirësohet ndjeshëm, nga vlera 0.6067 pa konsideruar metrikat e leximit deri në vlerën 0.6603 duke përfshirë në model veçoritë bazë të metrikave të leximit, duke sugjeruar se struktura e foljes shoqërohet me një ngarkesë më të lartë përpunuese gjatë leximit. Ngjashëm, rritje të konsiderueshme vërejmë edhe për kategoritë funksionale si CCONJ dhe ADP, si edhe për etiketat e ADJ, ADV dhe NUM. Në përgjithësi, rezultatet mbështesin hipotezën se integrimi i të dhënave nga gjurmimi i syrit përmirëson performancën e etiketimit të pjesëve të ligjëratës.

Tabela 6.13. Rezultatet e F1-score në eksperimentet për parashikimin e pjesëve të ligjëratës me arkitekturën BiLSTM duke përdorur vlerën e devijimit standard

POS Tag	**No Gaze**	**B**	**B+E**	**B+E+L**	**B+E+L+S**
NOUN	0.8575	0.8607	0.8408	**0.8624**	0.8580
PUNCT	0.9355	0.9343	0.9154	0.9189	**0.9673**
AUX	0.7889	0.7756	**0.8190**	0.8020	0.7960
DET	0.8417	0.8612	**0.8625**	0.8599	0.8617
PRON	0.7327	0.7744	0.7502	**0.7803**	0.7783
VERB	0.5322	0.6025	**0.6363**	0.6348	0.6553
CCONJ	0.8462	**0.9104**	0.8663	0.8801	0.8620
ADJ	0.6644	**0.6698**	0.6265	0.6418	0.6502
ADP	0.8247	0.8810	0.8627	**0.9088**	0.8840
PART	0.5704	**0.6163**	0.5765	0.5990	0.5958
ADV	0.5827	0.6416	0.6560	0.6713	**0.6742**
SCONJ	0.7598	0.7569	0.7521	**0.7861**	0.7608
PROPN	0.5038	0.6246	**0.6622**	0.5793	0.5886
NUM	0.6056	0.6029	0.5693	**0.6207**	0.5600

Heatmap-et e paraqitur në Figurën 6.9 dhe Figurën 6.10 tregojnë në mënyrë vizuale krahasimin e vlerave të F1-score për secilën etiketë në secilin konfigurim. Në boshtin vertikal paraqiten etiketat morfosintaksore, ndërsa në boshtin horizontal paraqiten shkurtimet për secilin nga konfigurimet eksperimentale.

Këto heatmap-e e bëjnë më të lehtë identifikimin e modeleve me performancën më të lartë për çdo etiketë, duke nxjerrë në pah kombinimet e parametrave që japin rezultatet më optimale. Vizualizimi i rezultateve ndihmon në vlerësimin e ndikimit të ndryshimeve të konfigurimeve eksperimentale mbi secilën etiketë morfosintaksore.

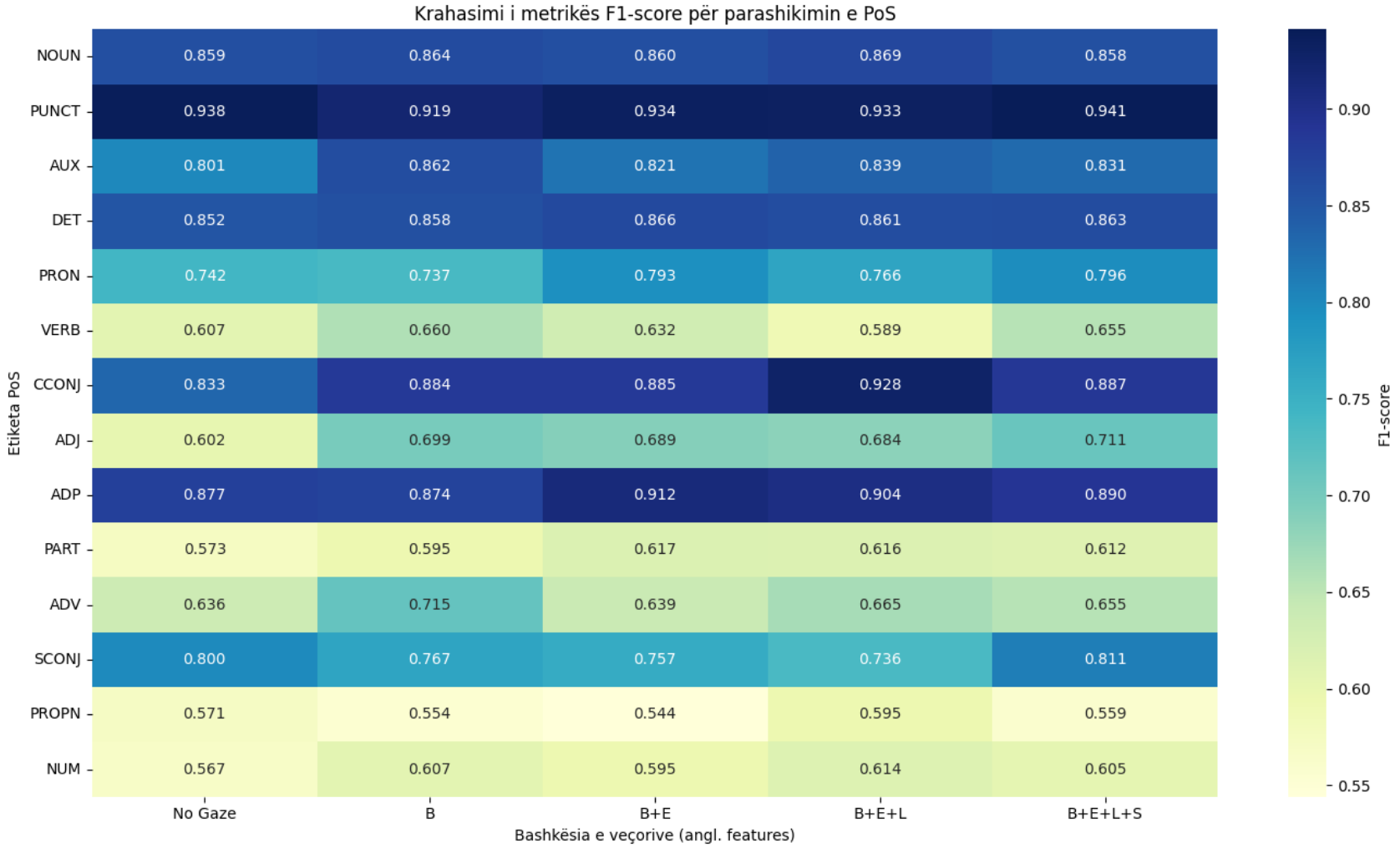


Figura 6.9. Heatmap i F1-score për parashikimin e pjesëve të ligjëratës me arkitekturën BiLSTM, bazuar në vlerat mesatare

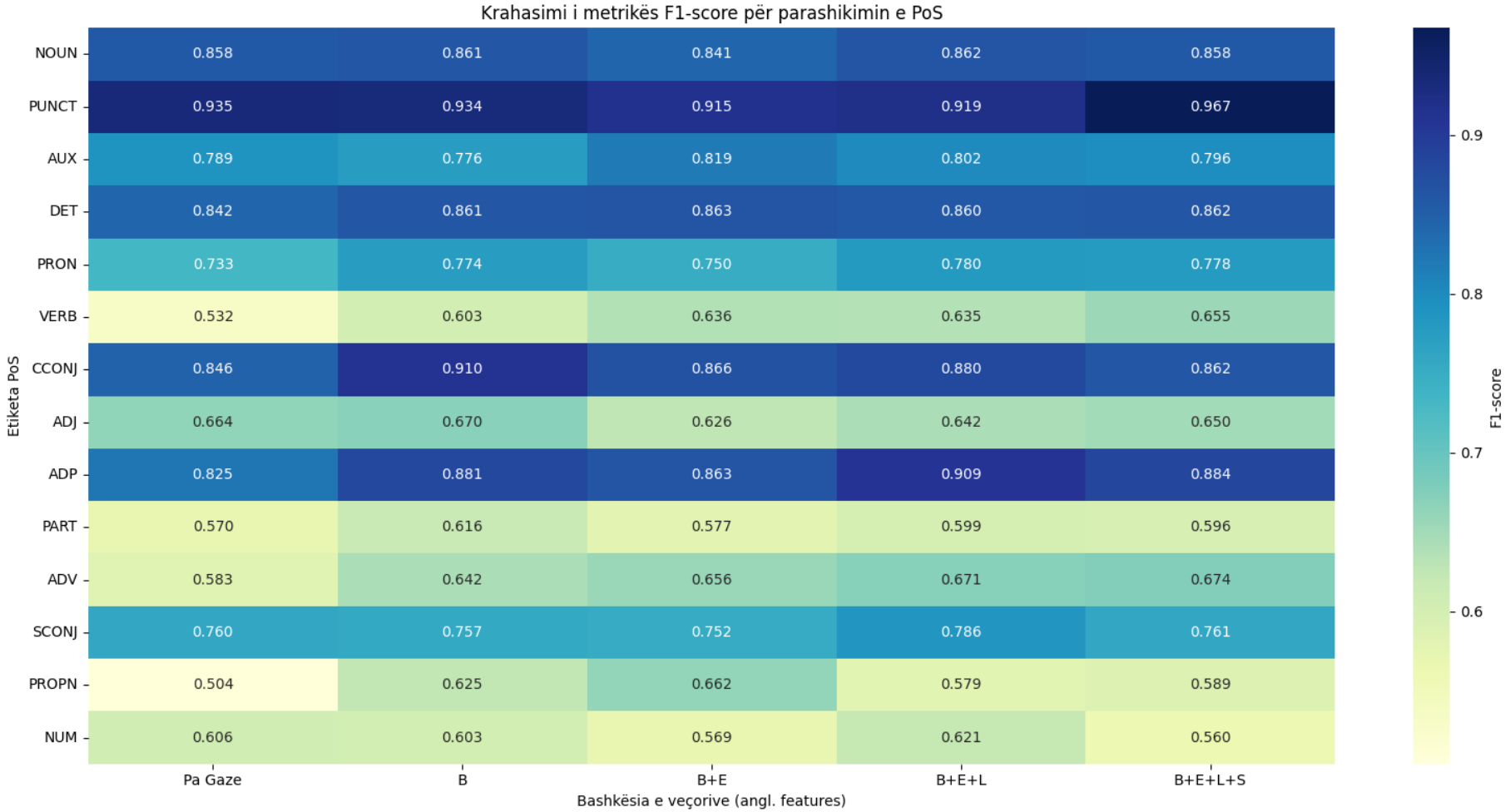


Figura 6.10. Heatmap i F1-score për parashikimin e pjesëve të ligjëratës me arkitekturën BiLSTM, bazuar në vlerat e devijimit standard

Në vijim të paraqitjes vizuale me heatmap, në Tabelën 6.14 paraqiten përmirësimet maksimale të arritura nga shtimi i metrikave të leximit për secilën etiketë të pjesëve të ligjëratës. Këtu mund të krahasohen më lehtë përmirësimet maksimale të arritura në dy rastet e studiuara, përkatësisht duke përdorur mesataren dhe devijimin standard të metrikave të leximit.

Krahasimi ndërmjet rezultateve të bazuara në vlerat mesatare (*mean*) dhe atyre të bazuara në devijimin standard (*std*) të metrikave të leximit tregon dallime në qëndrueshmërinë e modelit. Qasja e bazuar në devijimin standard të vlerave ofron një performancë më të qëndrueshme dhe metodologjikisht më të mirë për parashikimin e pjesëve të ligjëratës në krahasim me mesataren, duke marrë në konsideratë një rritje të metrikës weighted F1 nga 0.7678 në rastin e mesatares në 0.7969 në rastin e devijimit standard. Veçoritë e nxjerra nga variacioni ndër pjesëmarrës (*std*) duket se kapin më mirë dallimet në ngarkesën përpunuese të kategorive gramatikore sesa vlerat mesatare të leximit.

Tabela 6.14. Përmirësimet maksimale të F1-score gjatë eksperimenteve për parashikimin e pjesëve të ligjëratës me arkitekturën BiLSTM (vlerat mesatare)

ETIKETA POS	**NO GAZE**	**PËRMIRËSIMI MAKSIMAL (MEAN)**	**PËRMIRËSIMI MAKSIMAL (STD)**
NOUN	0.8592	**+0.0094**	+0.0049
PUNCT	0.9383	+0.0029	**+0.0318**
AUX	0.8014	**+0.0604**	+0.0301
DET	0.8522	+0.0137	**+0.0208**
PRON	0.7422	**+0.0538**	+0.0476
VERB	0.6067	+0.0536	**+0.1231**
CCONJ	0.8331	**+0.0945**	+0.0642
ADJ	0.6018	**+0.1089**	+0.0054
ADP	0.8775	+0.0344	**+0.0841**
PART	0.5734	+0.0437	**+0.0459**
ADV	0.6360	+0.0786	**+0.0915**
SCONJ	0.8000	+0.0107	**+0.0263**
PROPN	0.5705	+0.0243	**+0.1584**
NUM	0.5669	**+0.0474**	+0.0151

6.3.5. Ndërtimi dhe trajnimi i një modeli të bazuar në arkitekturën BiLSTM për parashikimin e lidhjeve sintaksore

Në këtë nënkapitull paraqitet procesi i trajnimit të një modeli BiLSTM për parashikimin e strukturës së varësive sintaksore. Qëllimi i modelit është të parashikojë për secilën fjalë në tekst etiketën HEAD dhe marrëdhënien e varësisë DEPREL. Ngjashëm me modelin në nënkapitullin 6.3.4, janë përdorur të dhënat e mbledhura nga eksperimentet e gjurmimit të syrit, ku përllogariten mesatarja dhe devijimi standard ndër pjesëmarrës të ndryshëm. Modeli kombinon informacionin leksikor, etiketat e pjesëve të ligjëratës (PoS) dhe metrikat e leximit nga eksperimentet e gjurmimit të syrit, duke ndjekur pesë konfigurimet e listuara në Tabelën 6.11. Gjithashtu konfigurimi i mjedisit të ekzekutimit dhe parametrave të modelit është i ngjashëm me modelin në nënkapitullin 6.3.4, me ndryshimin që në këtë rast numri i epokave të përdorura gjatë trajnimit është 20.

Tabelën 6.15 paraqet performancën e modelit BiLSTM me vlerat e UAS (*Unlabeled Attachment Score*) dhe LAS (*Labeled Attachment Score*) për identifikimin e marrëdhënieve sintaksore në pesë konfigurimet e konsideruara gjatë eksperimenteve duke përdorur vlerat mesatare të metrikave të leximit ndër pjesëmarrës. Rezultatet e paraqitura tregojnë se integrimi i informacionit nga eksperimentet e gjurmimit të syrit ndikon në përmirësimin e parashikimit të lidhjeve sintaksore. Metrika UAS mat përqindjen e fjalëve për të cilat modeli parashikon saktë kokën sintaksore, ndërsa metrika LAS mat saktësinë e kapjes së kokës dhe etiketës sintaksore. Në përgjithësi vërejmë se ndikimi i veçorive konjitive nuk është uniform për të gjitha marrëdhëniet sintaksore.

Një nga përmirësimet më të dukshme vërehet në marrëdhënien *obj* (*object*) ku modeli me veçori të plota nga shikimi arrin një vlerë UAS/LAS prej 0.8438/0.7031, krahasuar me vlerat 0.8125/0.5938 në modelin pa metrika leximi. Ky rezultat tregon se informacioni nga sjellja e leximit ndihmon modelin të identifikojë më saktë marrëdhëniet objekt-folje, që shpesh janë të ndërlidhura me integrimin e kontekstit dhe strukturës semantike. Një përmirësim i ngjashëm vihet re edhe për marrëdhëniet *nmod* dhe *advmod*, ku veçoritë e hershme të shikimit rrisin ndjeshëm performancën.

Një përmirësim shumë i madh vihet re për marrëdhënien *acl:relcl* ku performanca është shumë e ulët në modelin bazë (0.2000/0.2000), por përmirësohet me shtimin e metrikave në konfigurimin 'Base+Early' duke shkuar në vlerat 0.5333/0.5333. Ky përmirësim sugjeron se strukturat sintaksore më të gjata dhe komplekse përfitojnë nga integrimi i informacionit konjitiv.

Ndërkohë për marrëdhëniet funksionale më të thjeshta si *det, case* dhe *aux* vërejmë se performanca nuk përmirësohet duke shtuar informacionin nga përpunimi i lexuesve gjatë leximit. Kjo mund të arsyetohet me faktin se këto janë marrëdhënie sintaksore të rregullta dhe mund të parashikohen lehtë edhe vetëm duke përdorur informacionin leksikor dhe pjesët e ligjëratës.

Tabela 6.15. Rezultatet e modelit të analizës sintaksore me arkitekturën BiLSTM (vlerat mesatare)

		UAS/LAS				
Deprel	**N**	**No Gaze**	**B**	**B+E**	**B+E+L**	**B+E+L+S**
nsubj	99	0.7576/ 0.6364	0.6970/ 0.5556	0.7071/ 0.5758	0.7273/ 0.6061	0.7172/ 0.5455
cop	20	0.6000/ 0.5000	0.4500/ 0.4000	0.5000/ 0.4500	0.5000/ 0.4000	0.5000/ 0.4500
mark	123	0.7967/ 0.7317	0.7805/ 0.7073	0.7967/ 0.7073	0.7317/ 0.6748	**0.8049/ 0.7398**
root	63	0.7619/ 0.6825	**0.8095/ 0.7143**	0.7937/ 0.6825	0.6825/ 0.6190	0.7143/ 0.6825
case	132	0.9318/ 0.9091	0.9167/ 0.9091	0.8712/ 0.8636	0.9091/ 0.9015	0.9091/ 0.9015
obl	99	0.6667/ 0.5556	0.5960/ 0.4646	0.6162/ 0.5152	**0.7273/ 0.6061**	0.6061/ 0.4646
cc	54	0.7222/ 0.6667	0.6852/ 0.6481	**0.7407/ 0.6667**	**0.7407/ 0.6667**	0.7222/ 0.6667
conj	51	0.4118/ 0.3725	0.5098/ 0.4510	**0.5098/ 0.4510**	**0.5294/ 0.3922**	0.4706/ 0.3725
nummod	36	0.9167/ 0.6944	0.8056/ 0.6111	0.8056/ 0.6111	0.8889/ 0.6389	0.7500/ 0.5278
obj	64	0.8125/ 0.5938	**0.8125/ 0.6250**	**0.8281/ 0.6562**	0.8125/ 0.6562	**0.8438/ 0.7031**
det:poss	42	0.9286/ 0.4286	0.9762/ 0.6429	**1.0000/ 0.6905**	0.9524/ 0.6429	0.9762/ 0.3810
nmod	120	0.6750/ 0.6167	**0.6833/ 0.6250**	**0.7417/ 0.7083**	**0.7500/ 0.7000**	0.6917/ 0.6667
advmod	100	0.5600/ 0.4300	**0.6100/ 0.5100**	**0.6500/ 0.5800**	0.6200/ 0.5000	0.6400/ 0.5400
acl:relcl	15	0.2000/ 0.2000	**0.3333/ 0.2667**	**0.5333/ 0.5333**	0.4000/ 0.4000	**0.5333/ 0.4000**
advcl	21	0.6190/ 0.1429	0.5238/ 0.1905	0.3810/ 0.0952	**0.6190/ 0.2381**	0.2857/ 0.0476
det	176	0.9602/ 0.9261	**0.9716/ 0.9091**	0.9375/ 0.8693	0.9489/ 0.8523	0.9432/ 0.9091
amod	92	0.8370/ 0.7826	0.8370/ 0.7826	**0.8696/ 0.8261**	0.7391/ 0.7065	0.8152/ 0.7717
punct	140	0.6071/ 0.6071	**0.6429/ 0.6429**	**0.6571/ 0.6571**	0.6214/ 0.6214	0.6214/ 0.6214
aux	35	0.9429/ 0.8857	0.9143/ 0.8571	0.8857/ 0.7714	0.8571/ 0.8000	**0.8857/ 0.8000**

Si përfundim, analiza e rezultateve tregon se ndikimi i të dhënave konjitive varet nga lloji i marrëdhënies sintaksore. Marrëdhëniet që lidhen me integrimin semantik dhe struktura më komplekse të fjalisë, si *obj*, *nmod*, *advmod* dhe *acl:relcl* përfitojnë më shumë në krahasim me marrëdhëniet funksionale dhe më të rregullta nga ana sintaksore, si *det*, *case* dhe *aux*. Në veçanti, përdorimi i metrikave të leximit mund të ndihmojë modelet të fokusohen më mirë në elementët që kërkojnë përpunim konjitiv më të thellë. Kjo mbështet faktin se të dhënat e gjurmimit të syrit mund të ndihmojnë në përmirësimin e modelimit të strukturave sintaksore.

Figura 6.11 dhe Figura 6.12 tregojnë në mënyrë vizuale performancën e modeleve të parashikimit të varësive sintasore për secilin tip lidhje (*Deprel*) sipas konfigurimeve të veçorive (pa të dhëna nga eksperimentet e gjurmimit të syrit (No Gaze), B, B+E, B+E+L, B+E+L+S). Nga heatmap-et vërejmë qartë se disa marrëdhënie sintaksore përfitojnë dukshëm nga shtimi i veçorive nga gjurmimi i syrit, duke treguar përmirësime të matshme të metrikave UAS/ LAS. Ky vizualizim gjithashtu ndihmon në identifikimin e konfigurimeve më efektive dhe zgjedhjen e parametrave më optimalë për trajnime të ardhshme.

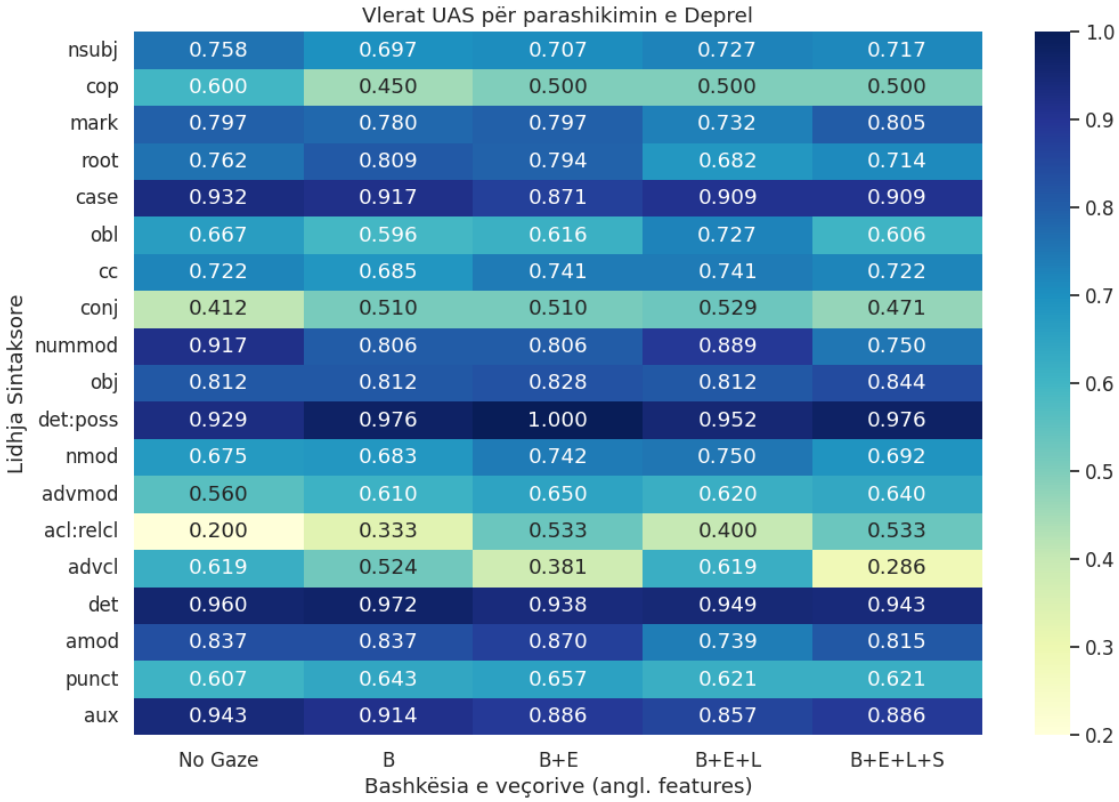


Figura 6.11. Heatmap i vlerës UAS për parashikimet e lidhjeve sintaksore me arkitekturën BiLSTM

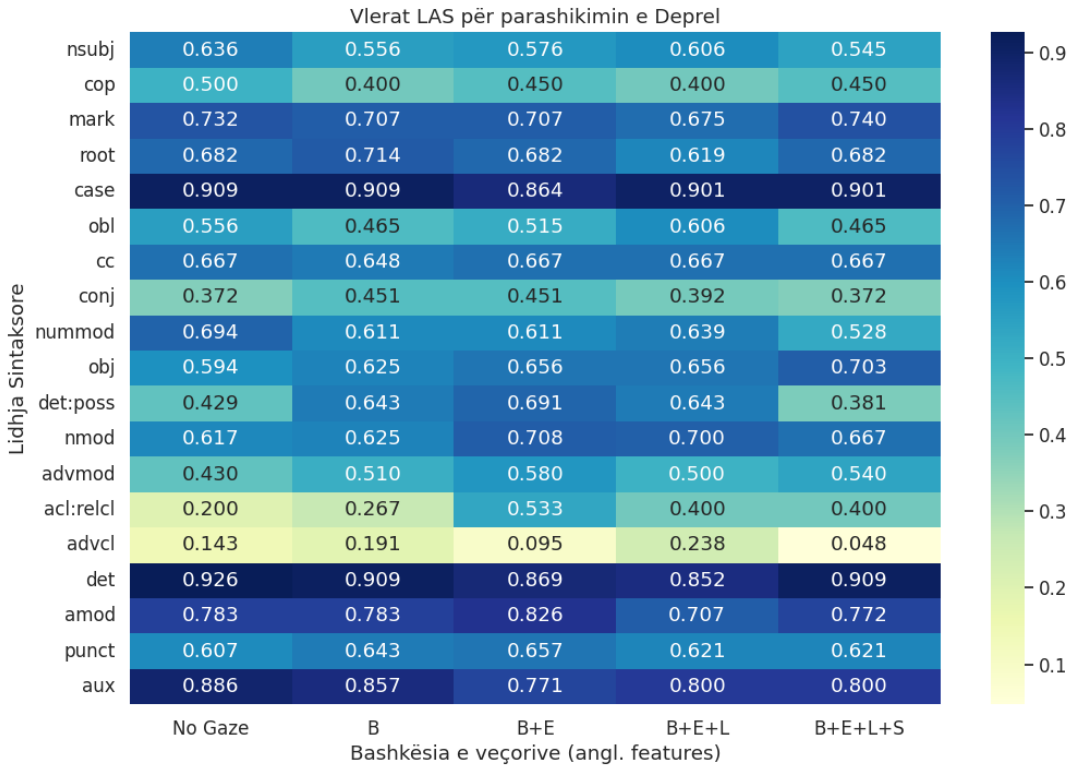


Figura 6.12. Heatmap i vlerës LAS për parashikimet e lidhjeve sintaksore me arkitekturën BiLSTM

6.4. Integrimi i veçorive gjuhësore dhe të dhënave konjitive për identifikimin e fjalëve kyçe

Nxjerrja e fjalë kyçeve (angl. *keyword extraction*) është një teknikë thelbësore për identifikimin e fjalëve ose frazave më të rëndësishme nga dokumentet në trajtë teksti, duke lehtësuar proceset e kërkimit, kategorizimit dhe përmbledhjes së informacionit. Në këtë seksion ne do të analizojmë dhe krahasojmë disa nga metodat kryesore të nxjerrjes së fjalëve kyçe për tekste në gjuhën shqipe duke i integruar ato me të dhënat konjitive të mbledhura nga eksperimentet e gjurmimit të syrit (Haveriku, Kote, Fone, & Meçe, 2024). Këto metoda përfshijnë teknika të bazuara në frekuencë, metoda statistikore, si dhe qasje të avancuara që përdorin modele të të mësuarit të makinës dhe mjete të përpunimit të gjuhës natyrore (PGJN-së).

Në mënyrë të veçantë, kemi eksploruar integrimin e modeleve YAKE (Campos, et al., 2020) dhe KeyBERT (Grootendorst, 2020) për gjuhën shqipe dhe përdorimin e metrikave të vlerësimit për të krahasuar saktësinë dhe përshtatshmërinë e këtyre metodave. Fjalët të cilat kërkojnë një kohë më të gjatë për t'u lexuar zakonisht janë fjalët më komplekse, të panjohura ose të rëndësishme nga ana semantike për tekstin e dhënë. Këto fjalë mund të identifikohen si fjalë kyçe, duke qenë se luajnë një rol të rëndësishëm në kuptimin e tekstit.

Gjatë përpunimit të të dhënave janë përllogaritur kohët mesatare të leximit për secilën fjalë. Pas llogaritjes së kohëve mesatare, të gjitha fjalët radhiten në rend zbritës në mënyrë që të kemi një pamje më të saktë të fjalëve të cilat janë vështruar më gjatë.

Gjatë eksperimenteve të kryera, qëllimi ynë ka qenë identifikohemi i fjalëve kyçe në tre prej teksteve: 'Hëna', 'Prilli i Thyer' dhe 'Njeriu i Shpellave', duke qenë se për këto tekste kemi informacion nga të dy tipet e eksperimenteve, ai i gjurmimit të mausit dhe të syrit.

Qëllimi kryesor i këtij implementimi është vlerësimi nëse informacioni i përfituar nga sjellja e leximit të përdoruesve mund të përdoret për të përmirësuar identifikimin dhe renditjen e fjalëve kyçe të gjeneruar nga algoritmat ekzistues, si edhe krahasimi midis rezultateve të arritura nga integrimi i të dhënave nga eksperimentet e gjurmimit të mausit dhe të syrit. Gjatë punës është përdorur një listë e fjalëve ndalëse (angl. *stopwords*) për gjuhën shqipe e cila përdoret për të pastruar fjalët funksionale, të cilat në rastin e eksperimenteve tona nuk përmbajnë informacion të dobishëm. Bashkësia e të dhënave të përdorura janë dy bashkësitë e paraqitura në KREU 4 (bashkësia nga eksperimentet e gjurmimit të mausit) dhe KREU 5 (bashkësia nga eksperimentet e gjurmimit të syrit), ku të dy bashkësitë janë të ndërlidhura edhe me veçoritë përkatëse gjuhësore. Të dhënat janë renditur sipas identifikuesit të tekstit, fjalisë dhe pozicionit të fjalës në mënyrë që mos të mos humbasim strukturën e tekstit. Si hap pastrimi është ndjekur eleminimi i shenjave të pikësimit duke filtruar etiketat të cilat përfshihet në shenjat e pikësimit (*PUNCT*), pasi nuk përmbajnë informacion të dobishëm për identifikimin e fjalëve kyçe.

Duke u bazuar në formatin e të dhënave në dispozicion kemi krijuar tre kategori kryesore:

1. Veçoritë e vëmendjes fillestare (angl. *attention*) që përfshijnë metrika si kohëzgjatja e fiksimit të parë dhe kohëzgjatjen e vështrimit.
2. Veçoritë e përpunimit konjitiv që lidhen me rikthimet dhe përpunimin e mëvonshëm të fjalëve.
3. Veçoritë e karakteristikave vizuale të leximit që përfshijnë numrin e fiksimeve dhe sakadave.

Të gjitha metrikat janë normalizuar duke përdorur transformimin *Min-Max scaling* në mënyrë që të dhënat të jenë të krahasueshme mes tyre. Formula e përzgjedhur për të përllogaritur pikëzimin konjitiv që do ti kalohet secilës fjalë është krijuar si një kombinim i ponderuar mes tre kategorive të listuara (0.5 për vëmendjen fillestare, 0.3 për përpunimin konjitiv dhe 0.2 për aspektet vizuale), me formulën e mëposhtme:

```
df["cognitive_score"] =
0.5*df["attention_score"] + 0.3*df["processing_score"]
+ 0.2*df["visual_score"]
```

Pikëzimi i përllogaritur nga kjo formulë synon të përfaqësojë rëndësinë relative të një fjale sipas përpunimit të lexuesve. Në Tabelën 6.16 paraqiten gjashtë konfigurimet e testuara gjatë eksperimenteve për identifikimin e fjalëve kyçe.

Tabela 6.16. Konfigurimet eksperimentale për identifikimin e fjalëve kyçe

KONFIGURIMI	**SHKURTIMI**
Algoritmi Yake bazë	YAKE Base
Algoritmi Yake me metrikat e leximit bazë + etiketën e pjesës së ligjëratës (*PoS*) + metrikat e leximit (*MOUSE*)	YAKE POS+MOUSE
Algoritmi Yake me metrikat e leximit bazë + etiketën e pjesës së ligjëratës (*POS*) + metrikat e leximit (*EYE*)	YAKE POS+EYE
Algoritmi KeyBERT bazë	KeyBERT Base
Algoritmi KeyBERT me metrikat e leximit bazë + etiketën e pjesës së ligjëratës (*PoS*) + metrikat e leximit (*MOUSE*)	KeyBERT POS+MOUSE
Algoritmi KeyBERT me metrikat e leximit bazë + etiketën e pjesës së ligjëratës (*PoS*) + metrikat e leximit (*EYE*)	KeyBERT POS+EYE

Në tabelë paraqiten dy algoritmet kryesore YAKE dhe KeyBERT, të cilët testohen në disa variante për të analizuar ndikimin e informacionit gjuhësor dhe konjitiv në saktësinë e rezultateve të paraqitura. Konfigurimet bazë përfaqësojnë versionin standard të algoritmeve, ku nxjerrja e fjalëve kyçe kryhet duke përdorur karakteristikat statistikore të tekstit. Ndërkohë, në konfigurimet e avancuara shtohen dy filtra të rinj, i pari është shtimi i etiketës së pjesëve të ligjëratës (*PoS*) i cili kufizon fjalët kandidate në varësi të kategorisë së tyre leksikore, dhe filtri i dytë që është shtimi i metrikave të leximit të përfituar nga gjurmimi i lëvizjeve të mausit (*MOUSE*) dhe nga gjurmimi i syrit (*EYE*). Metrikat e leximit përdoren për të modifikuar ose përmirësuar renditjen e fjalëve kyçe duke supozuar se fjalët që tërheqin më shumë vëmendje gjatë leximit mund të jenë më të rëndësishme për të përfaqësuar përmbajtjen e tekstit.

Gjatë përpunimit të teksteve kemi konsideruar lemën (angl. *lemma*) e fjalës në vend të formës që gjendet në tekst, për të reduktuar variacionin morfologjik të gjuhës shqipe. Gjithashtu, krijohet një variant i tekstit që përfshin vetëm fjalët me kategori morfosintaksore *NOUN*, *PROPN* dhe *ADJ*, pasi zakonisht këto janë kategoritë kryesore që mbartin përmbajtjen semantike të një teksti. Secila temë në secilin tekst lidhet me pikëzimin e saj konjitiv, të përllogaritur me formulën e mëparshme.

Procesi i rirenditjes kryhet përmes funksionit `rerank()` i cili merr si parametër fjalët kyçe (`phrase`) dhe pikëzimin bazë (`score`) të gjeneruar nga varianti bazë i algoritmit përkatës dhe pikëzimin nga formula e llogaritjes së informacionit konjitiv (`gaze`). Një parametër kontrolli (`gamma=0.7`) është përzgjedhur për të dhënë peshën e ndikimit të informacionit konjitiv në pikëzimin final.

```
def rerank(keywords, gaze, gamma=0.7):
    reranked = []
    for score, phrase in keywords:
        g = phrase_gaze_score(phrase, gaze)
        final = score * (1 + gamma * g)
        reranked.append((final, phrase))
    reranked.sort(reverse=True)
    return reranked
```

Për të inicializuar objektin KeyBERT, u përdorën parametrat e mëposhtëm:

1. Përdorimi i modelit *all-MiniLM-L6-v2* nga *Sentence Transformers* si model embedding.
2. Përcaktimi i parametrit *keyphrase_ngram_range* për 1-gram (unigrame). Synimi fillestar ishte të nxirreshin vetëm fjalë kyçe me një token.
3. Përcaktimi i parametrit $top_n = 10$ për të nxjerrë 10 fjalët kyçe kryesore nga teksti.

Për secilin tekst ruhen rezultatet për të gjithë konfigurimet e paraqitura në Tabelën 6.17. Rezultatet përfshijnë listën e fjalëve kyç të renditura sipas pikëzimit përkatës në secilin konfigurim. Rezultatet e paraqitura në tabelë tregojnë dallime të dukshme mes konfigurimeve të ndryshme të algoritmeve. Në momentin që në procesin e identifikimit të fjalëve kyçe përfshihen informacionet gjuhësore dhe të dhënat e ndërveprimit të përdoruesit (*PoS+Mouse* ose *PoS+Eye*) vihet re një ndryshim në renditjen dhe përzgjedhjen e fjalëve kyçe. Në konfigurimet e zgjeruara vihet re se favorizohen emrat dhe termat konceptuale që ka më shumë gjasa të përfaqësojnë përmbajtjen semantike të tekstit. Për shembull në tekstin "Hëna" konfigurimi *YAKE PoS+EYE* thekson termat si *'diametër'*, *'natyror'* dhe *'orbital'*, të cilat përfaqësojnë karakteristikat e lidhura me këtë objekt qiellor. Një identifikim i ngjashëm vihet re edhe në tekstin "Njeriu i Shpellave" ku termat '*njeri'*, '*ADN'*, '*bujqësi'* dhe '*sistem'* renditen në mënyrë më të qëndrueshme në krahasim me konfigurimet bazë.

Në konfigurimet me KeyBERT, shikojmë se integrimi i të dhënave konjtive ndikon kryesisht në renditjen e rëndësisë relative të termave ekzistues. Kjo sugjeron se integrimi i të dhënave shtesë gjuhësore dhe të dhënave konjitive (lëvizja e syrit apo e mausit) mund të përmirësojnë procesin e përzgjedhjes dhe renditjes së fjalëve kyçe, duke përmirësuar algoritmat tradicionalë të nxjerrjes së fjalëve kyçe në përpunimin e teksteve në gjuhën shqipe.

Më poshtë janë listuar fjalët kyçe të përzgjedhura manualisht për secilin nga tekstet e përfshira në eksperimente.

```
gold_keywords= {
1:
["veri","qytet","truall","pllajë","udhëtim","mal",
"rrafsh","shkrim","karrocë","Besian”]
2:
["njeri","shpellë","ADN","gjenom","shpikje","lëkurë",
"gjuetar","bujqësi","imunitar","evropian"]
3:
["hënë","tokë","satelit","diell","distancë","diametër",
"hënor","orbital","gravitacional","baticë"]
}
```

Këto fjalë kyçe janë përdorur për të përllogaritur vlerën e F1-score për secilin nga konfigurimet e përfshira në eksperiment. Analiza e rezultateve të paraqitura në Tabelën 6.18 tregon një ndikim të madh ndërmjet konfigurimeve të ndryshme.

Për shembull për tekstin “Prilli i Thyer” vihet re se algoritmi YAKE në konfigurimin bazë performon dobët ($F1-score=0.20$), shtimi i metrikave të mausit (PoS+MOUSE) e rrit vlerën ($F1-score=0.30$), ndërsa shtimi i metrikave nga leximi (PoS+EYE) përmirëson dukshëm rezultatin ($F1-score=0.50$). Nga ana tjetër algoritmi KeyBERT performon relativisht dobët në të gjitha konfigurimet.

Tabela 6.17. Lista e fjalëve kyçe të identifikuara në secilin konfigurim eksperimental

Teksti	**YAKE** **Bazë**	**YAKE** **PoS + MOUSE**	**YAKE** **PoS + EYE**	**KeyBERT** **Bazë**	**KeyBERT** **PoS + MOUSE**	**KeyBERT** **PoS + EYE**
Hëna	Tokë (124.99) periudhë (30.44) Diell (27.84) hënë (26.66) ndikim (22.78) baticë (19.46) zgjat (12.96) ditë (12.96) satelit (10.69) natyror (10.69)	Tokë (136.91) Diell (34.33) periudhë (33.66) hënë (31.30) baticë (29.35) ndikim (24.70) distancë (14.94) diametri(14.86) mbyllje (14.71) rrotullim (14.62)	Tokë (111.63) Hënë (42.45) Diell (36.22) periudhë (24.55) baticë (23.69) ndikim (18.94) diametër (11.14) orbital (11.11) natyror (11.05) anë (10.96)	gravitacional (0.2962) distancë (0.2898) orbital (0.2712) objekt (0.2661) satelit (0.2564) përballem (0.2543) barabartë (0.2490) rrotulloj (0.2415) km (0.2307) mbyllje (0.2229)	orbital (0.6806) distancë (0.6461) gravitacional (0.5447) satelit (0.5085) objekt (0.4244) diametri (0.3692) periudhë (0.3613) afërt (0.3526) hënor (0.3169) faktor (0.3111)	orbital (0.526) distancë (0.461) diametër (0.416) satelit (0.380) objekt (0.349) gravitacional (0.339) anë (0.272) periudhë (0.267) afërt (0.265) faktor (0.253)
Njeriu i shpellave	njeri (278.67) evropian (90.69) modernë (90.69) sistem (90.69) bujqësi (83.15) lëkurë (63.13) imunitar (63.13) blu (60.30) Evropë (59.38) Zbuloj (57.67)	njeri (102.01) evropian (68.49) sistem (67.07) modernë (63.19) bujqësi (59.70) Evropë (53.57) imunitar (53.11) lëkurë (43.50) Spanjë (43.17) blu (40.38)	njeri (79.76) ADN (63.51) sistem (50.84) Spanjë (47.41) evropian (46.00) bujqësi (45.28) Evropë (41.87) imunitar (41.39) blu (40.51) lëkurë (36.86)	shpikje (0.4836) pamje (0.4435) shfaqje (0.4352) mbetje (0.4273) përzierje (0.4054) përbërje (0.3975) pjekuri (0.3956) spanjë (0.3765) gjenom (0.3690) bujqësi (0.3667)	shpikje (0.5048) mbetje (0.4680) spanjë (0.4436) gjenom (0.3867) dietë (0.3736) shkencëtar (0.3650) njeri (0.3615) përbërje (0.3526) bujqësi (0.3477) bujqësor (0.3430)	mbetje (0.352) shpikje (0.337) spanjë (0.333) shfaqje (0.319) dietë (0.306) pjekuri (0.296) përzierje (0.296) përbërje (0.291) gjenom (0.285) pamje (0.282)
Prilli i thyer	veri (82.77) qytet (72.33) fundit (69.47) truall (69.47) akoma (69.47) them (69.47) shkrim (69.47) shkoj (69.47) Besian (65.23) Vorpsi (65.23)	Vorpsi (37.17) Besian (35.02) veri (25.93) Gumë (19.05) Kapitull (18.89) Dianën (18.84) THYER (18.20) Adrian (18.02) qytet (17.04) shkrim (16.99)	Besian (48.20) Vorpsi (33.21) veri (24.65) qytet (20.96) botë (20.21) udhëtim (20.09) vogël (19.87) mal (17.64) shkrim (17.57) dorë (17.43)	ngjethje (0.4203) shoqjë (0.3983) ngrij (0.3703) pamjë (0.3640) shpjegoj (0.3589) shkruaj (0.3550) shfaqje (0.3499) nisje (0.3413) ndjej (0.3313) shkoj (0.3301)	lumjë (0.5239) ngjethje (0.4989) shoqjë (0.4609) legjendë (0.4591) pabanuar (0.4567) shtojzovall (0.4405) shtojzovalle (0.4322) shkrimtar (0.4255) pamjë (0.4037) shfaqje (0.3936)	vogël (0.457) ngjethje (0.446) pllajë (0.429) qesëndisje (0.398) shfaqje (0.379) pështjellim (0.367) ndenjëse (0.347) pabanuar (0.332) mjaltë (0.326) çështje (0.309)

Në rastin e tekstit “Njeriu i Shpellave”, algoritmi YAKE ruan një performancë të qëndrueshme midis konfigurimeve. Ndërkohë, algoritmi KeyBERT përfiton nga shtimi i metrikave PoS+MOUSE. Në rastin e tekstit “Hëna”, të dy algoritmet arrijnë performancë të lartë, me rastin e YAKE që përfiton nga shtimi i metrikave të leximit dhe KeyBERT që ruan një vlerë të qëndrueshme por të lartë.

Tabela 6.18. Rezultat e F1-score për identifikimin e fjalëve kyçe sipas konfigurimeve eksperimentale

Teksti	**YAKE bazë**	**YAKE POS + EYE**	**YAKE POS + MOUSE**	**KeyBERT bazë**	**KeyBERT POS + EYE**	**KeyBERT POS + MOUSE**
Prilli i Thyer	0.20	**0.50**	0.30	**0.20**	0.10	0.00
Njeriu i shpellave	0.50	0.50	0.50	0.20	0.20	**0.40**
Hëna	0.50	**0.60**	0.50	0.40	**0.50**	**0.50**

Bazuar në rezultatet e paraqitura më lart arrijmë në përfundimin se algoritmi YAKE tregon një performancë më të qëndrueshme në shumicën e rasteve, veçanërisht kur shtohen metrikat e leximit, duke treguar se informacioni konjitiv ndihmon në identifikimin e fjalëve kyçe. Krahasimi i rezultateve tregon se efekti i përmirësimeve ndikohet nga natyra e tekstit, ku në rastin e teksteve narrative me emra vendesh, metrikat konjitive ndihmojnë në arritjen e rezultateve më të larta. Megjithatë, rezultatet e arritura tregojnë përfitimet e integrimit të informacionit konjitiv, duke theksuar nevojën për punime më të detajuara në të ardhmen.

Të dhënat konjitive nuk zëvendësojnë algoritmat ekzistues të nxjerrjes së fjalëve kyçe, por funksionojnë si shtresë plotësuese për përmirësimin e renditjes dhe filtrimit semantik.

KREU 7

PËRFUNDIME

Në këtë disertacion është studiuar ndërthurja e dy fushave: studimet psikolinguistike të leximit dhe përpunimi i gjuhës natyrore, duke u përqendruar te gjuha shqipe. Qëllimi kryesor ka qenë të analizohet nëse të dhënat e mbledhura gjatë eksperimenteve të gjurmimit të syrit dhe të mausit mund të përdoren si informacion shtesë për të përmirësuar disa detyra të ndërlidhura me përpunimin e tekstit në gjuhën shqipe. Nisur nga fakti që shqipja është një gjuhë me burime të kufizuara në fushën e përpunimit automatik të gjuhës, ky studim ndërtohet mbi një problem praktik: mungesa e korpuseve të mëdha dhe burimeve të hapura. Duke qenë se krijimi i korpuseve të mëdha kërkon një angazhim të konsiderueshëm nga ekspertë gjuhësorë, qasjet e reja që mund të realizohen në një kohë më të shkurtër dhe me një popullatë më të gjerë paraqiten si një mundësi e vlefshme për të ofruar metoda alternative për përpunimin e teksteve në gjuhën shqipe. Gjithashtu, studimi i këtyre metodologjive të reja në një gjuhë si shqipja ka potencial për t'u aplikuar në gjuhë e tjera me burime të kufizuara.

Puna kërkimore nuk është ndalur vetëm te analiza teorike e mundësive që japin të dhënat konjitive, por është përqendruar në ndërtimin e bashkësive të reja të të dhënave për shqipen, në përcaktimin e një metodologjie për përpunimin e këtyre të dhënave dhe në përdorimin e tyre në eksperimente konkrete. Në këtë mënyrë, ky disertacion mbështetet në një qasje të dyfishtë: nga njëra anë ndërtimi i burimeve të reja dhe, nga ana tjetër, vlerësimi i përdorimit të tyre në detyra konkrete të përpunimit të tekstit.

Një nga rezultatet më të rëndësishme të këtij disertacioni është krijimi i dy bashkësive të para të të dhënave të këtij lloji për gjuhën shqipe. E para është bashkësia e të dhënave nga gjurmimi i kursorit të mausit, e ndërtuar duke zhvilluar eksperimente me 51 pjesëmarrës, folës nativë të gjuhës shqipe. E dyta, është bashkësia e të dhënave nga gjurmimi i syrit, e ndërtuar me pajisjen EyeLink Portable Duo, me 25 pjesëmarrës, folës nativë të gjuhës shqipe. Këto dy bashkësi përbëjnë një kontribut të drejtpërdrejtë për shqipen, sepse krijojnë për herë të parë një bashkësi të dhënash që mund të përdoren për të studiuar mënyrën se si lexuesit përpunojnë tekstin në shqip.

Është e rëndësishme të theksojmë se për të dyja bashkësitë e të dhënave janë përcaktuar dhe dokumentuar të gjithë hapat, duke nisur nga arsyeja e përzgjedhjes së teksteve, veçoritë teknike të implementimit të eksperimenteve dhe deri te parapërpunimi i të dhënave për gjenerimin e metrikave të leximit. Ky dokumentim i hollësishëm lehtëson riprodhueshmërinë e eksperimenteve nga studiues të tjerë që synojnë zgjerimin e këtyre bashkësive të të dhënave për gjuhën shqipe apo të zhvillojnë eksperimente të ngjashme për gjuhë të tjera me burime të kufizuara. Përtej krijimit të bashkësive të të dhënave është realizuar edhe një analizë e sjelljes së pjesëmarrësve gjatë leximit në varësi të moshës. Në rastin e eksperimenteve të gjurmimit të syrit, kjo analizë është zgjeruar më tej duke përfshirë informacionin e mbledhur nga pyetësorët e pjesëmarrësve. Konkretisht janë analizuar faktorë si kohëzgjatja e eksperimentit në

varësi të moshës, individëve, gjinisë dhe nivelit të arsimit, duke ofruar një pasqyrë më të detajuar të variacioneve të sjelljes gjatë leximit.

Eksperimentimi me dy metodologji eksperimentale ka një qëllim më të gjerë: të vlerësojë nëse alternativat me kosto të ulët, si gjurmimi i kursorit të mausit, mund të ofrojnë rezultate të krahasueshme me alternativat më të shtrenjta, si gjurmimi i syrit përmes pajisjes EyeLink Portable Duo. Konkretisht, nga analiza e metrikave të leximit në nivel fjale, duke krahasuar bashkësitë e të dhënave të krijuara nga eksperimentet e gjurmimit të syrit dhe të mausit, në kreun 5.6, u vërtetua ekzistenca e një korrelacioni pozitiv, veçanërisht për metrikën e kohëzgjatjes totale të leximit për çdo fjalë. Rezultatet e paraqitura tregojnë se teknika e gjurmimit të mausit mund të shërbejë si një alternativë me kosto të ulët, veçanërisht për gjuhët që nuk kanë akses në laboratorë me pajisje më të avancuara për gjurmimin e syrit.

Një nga sfidat e identifikuara nga rishikimi i literaturës ishte nevoja e zgjerimit të të dhënave konjitive gjatë leximit në shumë gjuhë, në mënyrë që të studiohet në mënyrë krahasimore integrimi i këtyre të dhënave në modelet e inteligjencës artificiale për detyra të ndryshme në fushën e PGJN-së. Në rastin e gjuhës shqipe, mungesa e burimeve paraqet sfida për zhvillimin e modeleve të etiketimit automatik. Përdorimi i sinjaleve konjitive ofron një burim alternativ që mund të kompensojë mungesën e të dhënave gjuhësore. Të dhënat konjitive mund të ndihmojnë në modelet ekzistuese për të mësuar ndarje më të imta mes kategorive gramatikore ose për të ndihmuar në aftësitë përgjithësuese të modeleve.

Në këtë studim integrimi i të dhënave konjitive u testua në tri detyra kryesore: parashikimi i pjesëve të ligjëratës, parashikimi i lidhjeve sintaksore dhe identifikimi i fjalëve kyçe. Për të dyja bashkësitë e të dhënave u paraqit në trajtë grafike shpërndarja e kohëzgjatjes së vështrimit sipas etiketave të pjesëve të ligjëratës dhe marrëdhënieve sintaksore të varësisë. Rezultatet e arritura nga analizimi i kësaj shpërndarje vërtetuan se për gjuhën shqipe, ngjashëm me gjuhën angleze, përqendrimi më i madh gjatë leximit qëndron në fjalët e përmbajtjes (etiketat PoS: *NOUN*, *VERB*, *ADJ*, *ADV*). Për sa i përket marrëdhënieve sintaksore të varësisë, elementët e strukturës së fjalisë (etiketat deprel: *nsubj*, *obj*, *ccomp*, *xcomp*) shfaqin medianë më të lartë të kohëzgjatjes së leximit në krahasim me marrëdhëniet më periferike dhe funksionale. Në këtë mënyrë konfirmojmë se leximi në gjuhën shqipe ka modele të ngjashme të vëmendjes vizuale me gjuhën angleze, duke motivuar zhvillimin e studimeve krahasuese në të ardhmen.

Duke marrë në konsideratë modelet më të përdorura në punime të kësaj fushe, si edhe duke u bazuar në veçoritë e përzgjedhura dhe madhësinë e bashkësive tona të të dhënave kemi eksperimentuar me një model statistikor CRF-Conditional Random Fields dhe me një arkitekturë më të avancuar BiLSTM për të trajnuar modelet për parashikimin e pjesëve të ligjëratës. Modeli CRF u testua për të dhënat nga gjurmimi i syrit dhe i mausit, duke arritur një performancë më të mirë në rastin e parë duke qenë se bashkësia e të dhënave nga gjurmimi i syrit është më e larmishme. Ngjashëm me punimet në gjuhë të tjera, të identifikuara gjatë rishikimit të literaturës, edhe për gjuhën shqipe vërejmë një rritje të metrikës së vlerësimit të modelit (F1-score) me vlera +0.1 deri në +0.3, duke krahasuar konfigurimin pa përfshirë dhe duke përfshirë metrikat e leximit.

Në rastin e arkitekturës BiLSTM u përdorën pesë konfigurime eksperimentale: pa metrika leximi (NG), me metrikat e leximit bazë (B), me metrika bazë dhe të herëta (B+E), me metrika bazë, të herëta dhe të vona (B+E+L) dhe konfigurimi i fundit me metrikat bazë, të herëta, të vona dhe të dhënat nga sakadat e lëvizjes së vështrimit (B+E+L+S). Arkitektura BiLSTM është përdorur për të eksperimentuar vetëm me të dhënat nga bashkësia e gjurmimit të syrit, për shkak se bashkësia e gjurmimit të mausit është më pak e larmishme në tekstet dhe fjalët e përfshira. Në rastin e parashikimit të pjesëve të ligjëratës është arritur një rritje e metrikës së F1-score, me përmirësime në etiketat si *CCONJ* (+0.09), *ADV* (+0.09), *ADP* (+0.08), *AUX* (+0.06) dhe *VERB* (+0.06). Rezultatet e parashikimit të etiketave të marrëdhënieve sintaksore u vlerësuan përmes metrikave UAS dhe LAS, ku u vërejt një përmirësim i performancës së modelit me integrimin e metrikave të leximit, sidomos në etiketat: det:poss (UAS: 1.0000/LAS: 0.6905), det (UAS: 0.9716/LAS: 0.9091), case (UAS: 0.9318/LAS: 0.9091) dhe aux (UAS: 0.9429/LAS: 0.8857).

Në detyrën e nxjerrjes së fjalëve kyçe është analizuar integrimi i metrikave të leximit nga eksperimentet e gjurmimit të syrit dhe mausit në modelet tradicionale YAKE dhe KeyBERT. Qëllimi kryesor i këtij implementimi është vlerësimi nëse informacioni i përfituar nga sjellja e leximit të përdoruesve mund të përdoret për të përmirësuar identifikimin dhe renditjen e fjalëve kyçe të gjeneruar nga algoritmat ekzistues, si edhe krahasimi midis rezultateve të arritura nga integrimi i të dhënave nga eksperimentet e gjurmimit të mausit dhe syrit. Janë eksperimentuar gjashtë konfigurime eksperimentale për identifikimin e fjalëve kyçe duke përfshirë algoritmin YAKE/KeyBERT bazë, me metrikat e leximit nga gjurmimi i mausit dhe nga gjurmimi i syrit. Bazuar në rezultatet e paraqitura në punim arrijmë në përfundimin se algoritmi YAKE tregon një performancë më të qëndrushme në shumicën e rasteve, veçanërisht kur shtohen metrikat e leximit, duke treguar se informacioni konjitiv ndihmon në identifikimin e fjalëve kyçe. Krahasimi i rezultateve tregon se efekti i përmirësimeve ndikohet nga natyra e tekstit.

Ky studim dhe rezultatet e arritura mbështesin hipotezën se të dhënat konjitive nga përpunimi njerëzor i gjuhës janë një burim i vlefshëm për t'u përdorur në ndërtimin apo përmirësimin e modeleve inteligjente kompjuterike. Gjithashtu, burimet e të dhënave të krijuara ndihmojnë në përfaqësimin e gjuhës shqipe në kërkimin ndërkombëtar, duke e bërë shqipen të krahasueshme me gjuhë të tjera.

Megjithatë, mbetet akoma punë për të ardhmen. Një nga pikat kryesore ku mund të përqendrohet vijueshmëria e punës është zgjerimi i bashkësive të të dhënave konjitive me më shumë pjesëmarrës duke krijuar një bashkësi më të larmishme, si edhe zgjerimi i tipeve të teksteve të përfshira në eksperimente për të rritur variacionin stilistik. Një tjetër pikë interesante për t'u studiuar do ishte zhvillimi i eksperimenteve me grupmosha të ndryshme dhe me folës të dialekteve të ndryshme për të analizuar ndryshimet individuale në sjelljen gjatë leximit sipas faktorëve të ndryshëm demografik.

Më tej me vlerë është krahasimi i metrikave të leximit mes bashkësisë së të dhënave në gjuhën shqipe dhe në bashkësi të tjera të dhënash paralele në gjuhë të tjera, e cila në këtë studim nuk është zhvilluar pasi krahasimi mund të zhvillohet midis të dhënave të

cilat janë zhvilluar në të njëjta kushte eksperimentale. Me zgjerimin e bashkësive të të dhënave mund të testohet integrimi i metrikave konjtive në arkitektura më të avancuara si p.sh. modelet e mëdha gjuhësore (*MMGJ*). Përdorimi i bashkësisë së të dhënave të mbledhura nga testet psikometrike mund të përdoret më tej për të identifikuar lidhjen mes memories përpunuese dhe sjelljes individuale gjatë leximit. Ndërkohë, nënfusha të tjera të PGJN-së ku mund të integrohen të dhënat konjitive janë njohja e entiteteve emërore, zbulimi i gabimeve gramatikore dhe identifikimi i shprehjeve shumëfjalëshe (angl. *multiword expressions*). Përdorimi i metrikave konjitive mund të analizohet gjithashtu për identifikimin automatik të kompleksitetit gjuhësor dhe ndërtimin e treguesve të vështirësisë së tekstit në gjuhën shqipe.

Në përfundim ky punim vendos një bazë për integrimin e të dhënave kompjuterike në disa detyra të përpunimit të teksteve në gjuhën shqipe. Në këtë mënyrë hapet rruga për zhvillime të mëtejshme në ndërtimin e modeleve më të avancuara, duke përafruar akoma më shumë modelet kompjuterike me mënyrën njerëzore të përpunimit të gjuhës.

BIBLIOGRAFIA

[1] Agalliu, F., Angoni, E., Demiraj, S., Dhrimo, A., Hysa, E., Lafe, E., & Likaj, E. (2002). *Gramatika e Gjuhës Shqipe.* Tiranë: Instituti i Gjuhësisë dhe i Letërsisë, Akademia e Shkencave e Shqipërisë.

[2] Akademia e Shkencave e Shqipërisë. (1980, 2002, 2006). *Fjalori i gjuhës së sotme shqipe.* (ASHSH).

[3] Anwyl-Irvine, A. L., Armstrong, T., & Dalmaijer, E. S. (2022). MouseView.js: Reliable and valid attention tracking in web-based experiments using a cursor-directed aperture. *Behavior Research Methods, 54*, 1663-1687. doi:https://doi.org/10.3758/s13428-021-01703-5

[4] Augereau, O., Jackuet, C., Kise, K., & Journet, N. (2018). Vocabulometer: A Web Platform for Document and Reader Mutual Analysis. *13th IAPR International Workshop on Document Analysis Systems* (fv. 109-114). Vienna, Austria: IEEE. doi:10.1109/DAS.2018.59

[5] Aunsri, N., & Rattarom, S. (2022). Novel eye-based features for head pose-free gaze estimation with web. *Ain Shams Engineering Journal, 13*(5). doi:https://doi.org/10.1016/j.asej.2022.101731

[6] Barret, M., Bingel, J., Hollenstein, N., Rei, M., & Søgaard, A. (2018). Sequence classification with human attention. *Proceedings ofthe 22nd Conference on Computational Natural Language Learning (CoNLL 2018).* Brussels, belgium: Association for Computational Linguistics. doi:10.18653/v1/K18-1030

[7] Barret, M., Keller, F., & Sogaard, A. (2016). Cross-lingual Transfer of Correlations between Parts of Speech and Gaze Features. *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers* (fv. 1330–1339). Osaka, Japan: The COLING 2016 Organizing Committee. Gjetur në https://aclanthology.org/C16-1126

[8] Barrett, M., & Hollenstein, N. (2020). Sequence labelling and sequence classification with gaze: Novel uses of eye-tracking data for Natural Language Processing. *Language and Linguistics Compass, 14*(11), 1-16. doi:https://doi.org/10.1111/lnc3.12396

[9] Barrett, M., & Søgaard, A. (2015). Reading behavior predicts syntactic categories. *Proceedings of the 19th Conference on Computational Language Learning* (fv. 345-349). Beijing, China: Association for Computational Linguistics. doi:10.18653/v1/k15-1038

[10] Barrett, M., Bingel, J., Keller, F., & Søgaard, A. (2016). Weakly Supervised Part-of-speech Tagging Using Eye-tracking Data. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)* (fv. 579–584). Berlin, Germany: Association for Computational Linguistics. doi:10.18653/v1/P16-2094

[11] Bautista, L. G., & Naval, P. J. (2020). Towards Learning to Read Like Humans. *12th International Conference on Computational Collective Intelligence. 12496*, fv. 779-791. Springer Science and Business Media Deutschland GmbH. doi:10.1007/978-3-030-63007-2_61

[12] Behseta, S., Berdyyeva, T., Olson, C. R., & Kass, R. E. (2009, Jan). Bayesian Correction for Attenuation of Correlation in Multi-Trial Spike Count Data. *J Neurophysiol.*, 2186-2193. doi:10.1152/jn.90727.2008

[13] Beinborn, L., & Hollenstein, N. (2024). *Cognitive Plausibility in Natural Language Processing* (bot. i Synthesis Lectures on Human Language Technologies). Springer Nature Switzerland AG. doi:https://doi.org/10.1007/978-3-031-43260-6

[14] Brandt, S., & Hollenstein, N. (2022). Every word counts: A multilingual analysis of individual human alignment with model attention. *Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)* (fv. 72-77). Association for Computational Linguistics. doi:10.18653/v1/2022.aacl-short.10

[15] Campos, R., Mangaravite, V., Parsquali, A., Jorge, A., Nunes, C., & Jatowt, A. (2020). YAKE! Keyword extraction from single documents using multiple local features. *509*(C), 257-289. doi:https://doi.org/10.1016/j.ins.2019.09.013

[16] Carter, B. T., & Luke, S. G. (2020). Best practices in eye tracking research. *International Journal of Psychophysiology, 155*, 49-62. doi:https://doi.org/10.1016/j.ijpsycho.2020.05.010

[17] Changwani, A., & Sarode, T. (2019). Low-cost Eye-Tracking for Foveated Rendering Using Machine Learning. *IEEE International Conference on Cloud Computing in Emerging Markets (CCEM)* (fv. 32-39). Bengaluru, India: IEEE. doi:10.1109/CCEM48484.2019.000-1

[18] Chen, X., Mao, J., Liu, Y., Zhang, M., & Ma, S. (2022). Investigating human reading behavior during sentiment judgment. *International Journal of Machine Learning and Cybernetics*, 2283–2296. doi:https://doi.org/10.1007/s13042-022-01523-9

[19] Cheng, S., Ping, Q., Wang, J., & Chen, Y. (2022). EasyGaze: Hybrid eye tracking approach for handheld. *Virtual Reality & Intelligent Hardware, 4*(2), 173-188. doi:https://doi.org/10.1016/j.vrih.2021.10.003

[20] Cop, U., Dirix, N., Drieghe, D., & Duyck, W. (2017). Presenting GECO: An eyetracking corpus of monolingual and bilingual sentence reading. *Behavior Research Methods volume 49*, fv. 602-615. doi: https://doi.org/10.3758/s13428-016-0734-0

[21] Denckla, M. B., & Cutting, L. E. (1999). History and significance of rapid automatized naming. *Part I Rapid Serial Naming And The Double-Deficit Hypothesis Ann. of Dyslexia, 49*, 29-42. doi:https://doi.org/10.1007/s11881-999-0018-9

[22] Duchowksi, A. T. (2017). *Eye Tracking Methodology* (bot. i 3). Springer Nature Switzerland AG 2017: Springer Cham. doi:https://doi.org/10.1007/978-3-319-57883-5

[23] Ebert, C., Islamaj, A., Kuqi, A., Sonnenhauser, B., Plamada, M., & Widmer, P. (2022). UD Gheg Pear Stories: An annotated treebank of Gheg Albanian as spoken in Switzerland. doi:10.21203/rs.3.rs-2056973/v1

[24] Eide, M. G., Watanabe, R., Heldal, I., Helgesen, C., Geitung, A., & Soleim, H. (2019). Detecting oculomotor problems using eye tracking: Comparing EyeX and TX300. *10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)* (fv. 381-388). Naples, Italy: IEEE. doi:10.1109/CogInfoCom47531.2019.9089895

[25] Engbert, R., Trunkenbrod, H. A., Barthelme, S., & Wichman, F. A. (2015). *Spatial statistics and attentional dynamics in scene viewing* (Vëll. i 15). Journal of Vision. doi:https://doi.org/10.1167/15.1.14

[26] Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*, 143-149. doi:https://doi.org/10.3758/BF03203267

[27] Frank, S. L., & Aumeistere, A. (2023). An eye-tracking-with-EEG coregistration corpus of narrative sentences. *Language Resources & Evaluation*. doi:https://doi.org/10.1007/s10579-023-09684-x

[28] Golhahn, D., Eckart, T., & Quasthoff, U. (2012). Building Large Monolingual Dictionaries at the Leipzig Corpora Collection: From 100 to 200 Languages. *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, fv. 759–765. Istanbul, Turkey: European Language Resources Association (ELRA). Gjetur në https://aclanthology.org/L12-1154/

[29] Gomez-Poveda, J., & Gaudioso, E. (2016). Evaluation of temporal stability of eye tracking algorithms using webcams. *Expert Systems with Applications, 64*, 69-83. doi:https://doi.org/10.1016/j.eswa.2016.07.029

[30] Grootendorst, M. (2020). KeyBERT: Minimal keyword extraction with BERT. doi:https://doi.org/10.5281/zenodo.4461265

[31] Guibon, G., Courtin, M., Gerdes, K., & Guillaume, B. (2020). When collaborative treebank curation meets graph grammars. *Proceedings of the Twelfth Language Resources and Evaluation Conference* (fv. 5291-5300). Marseille, France: European Language Resources Association. Gjetur në https://aclanthology.org/2020.lrec-1.651/

[32] Gunawardena, N., Ginige, J. A., Javadi, B., & Lui, G. (2022). Performance Analysis of CNN Models for Mobile Device Eye. *26th International Conference on Knowledge-Based and Intelligent Information & Engineering (KES 2022), 207*, 2291-2300. doi:https://doi.org/10.1016/j.procs.2022.09.288

[33] Haveriku, A., & Frashëri, N. (2023). Twitter Sentiment Analysis for Albanian Language. *The Albanian Journal of Natural and Technical Sciences, 1*. Gjetur në https://akad.gov.al/albanian-journal-of-natural-and-technical-sciences-ajnts-20231-vol-xxviii-57/

[34] Haveriku, A., & Meçe, E. K. (2025). Albanian Sentiment Analysis: Exploring Machine and Deep Learning Techniques. *Albanian Journal of Natural and Technical Sciences, Vol. XXIX (62)*. Gjetur në https://akad.gov.al/albanian-journal-of-natural-and-technical-sciences-ajnts-2025-1-vol-xxix62/

[35] Haveriku, A., Bedulla, S., & Meçe, E. K. (2025). Understanding Reading Patterns of Albanian Native Readers through Mouse Tracking Analysis. *The 39th International Conference on Advanced Information Networking and Applications (AINA-2025). 245.* Springer, Cham. doi:https://doi.org/10.1007/978-3-031-87763-6_38

[36] Haveriku, A., Haxhija, I., & Meçe, E. K. (2024). Analyzing Reading Patterns with Webcams: An Eye-Tracking Study of the Albanian Language. *International Conference on Software, Telecommunications and Computer Networks (SoftCOM)* (fv. 1-5). Split, Croatia: IEEE. doi:10.23919/SoftCOM62040.2024.10721886

[37] Haveriku, A., Kote, N., & Meçe, E. K. (2025). Comparing Word Level Reading Metrics From Eye-Tracking and Mouse-Tracking Approaches. *ICITEE25: 3nd International Conference on Information Technologies and Educational Engineering.* Tirana, Albania: ISBN: 9789928481474.

[38] Haveriku, A., Kote, N., Çepani, A., Çerpja, A., Trandafili, E., & Meçe, E. K. (20225). Leveraging mouse tracking data for part-of-speech and syntactic dependency prediction in Albanian. *ETRA '25: Proceedings of the 2025 Symposium on Eye Tracking Research and Applications* (fv. 1-5). New York, NY, USA: Association for Computing Machinery. doi:10.1145/3715669.3726840

[39] Haveriku, A., Kote, N., Fone, F., & Meçe, E. K. (2024). An exploration of keyword extraction approaches and eye-tracking solutions for the Albanian language. (ISBN: 9789928809636).

[40] Haveriku, A., Paci, H., Kote, N., & Meçe, E. K. (2024). Systematic Review of Eye-Tracking Studies. Në L. Barolli (Re.), *International Conference on Emerging Internet, Data & Web Technologies EIDWT 2024, Lecture Notes on Data Engineering and Communications Technologies. 193.* Springer, Cham. doi:https://doi.org/10.1007/978-3-031-53555-0_24

[41] Haveriku, A., Paci, H., Kote, N., Shasivari, P., & Meçe, E. K. (2025). A systematic review of eye-tracking data in NLP: exploring low-cost and cross-lingual possibilities. *International Journal of Grid and Utility Computing*, 29-40. doi:https://doi.org/10.1504/IJGUC.2025.143878

[42] Heinecke, J. (2019). ConlluEditor: a fully graphical editor for universal dependencies treebank files. *Proceedings of the Third Workshop on Universal Dependencies (UDW SyntaxFest 2019)* (fv. 87-93). Paris, France: Association for Computational Linguistics. doi:10.18653/v1/W19-8010

[43] Hollenstein, N., & Zhang, C. (2019). Entity Recognition at First Sight: Improving NER with Eye Movement Information. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (fv. 1–10). Minneapolis, Minnesota: Association for Computational Linguistics. doi:10.18653/v1/N19-1001

[44] Hollenstein, N., Barrett, M., & Bjornsdottir, M. (2022). The Copenhagen Corpus of Eye Tracking Recordings from Natural. *Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022)*, (fv. 1712-1720). Marseille. doi:https://aclanthology.org/2022.lrec-1.182.pdf

[45] Hollenstein, N., Pirovano, F., Zhang, C., Beinborn, L., & Jäger, L. (2021). Multilingual Language models Predict Human Reading Behavior. *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (fv. 106–123). Online: Association for Computational Linguistics. doi:10.18653/v1/2021.naacl-main.10

[46] Hollenstein, N., Rotsztejn, J., Troendle, M., Pedroni, A., Zhang, C., & Langer, N. (2018). Data Descriptor: ZuCo, a simultaneous EEG and eye-tracking resource for natural sentence reading. *Scientific Data volume 5*. doi:https://doi.org/10.1038/sdata.2018.291

[47] Honnibal, M., Montani, I., Landeghem, S. V., & Boyd, A. (2020). *spaCy: Industrial-strength Natural Language Processing in Python.* Zenodo. doi:10.5281/zenodo.1212303

[48] Hosp, B., Eivazi, S., Maurer, M., Fuhl, W., Geisler, D., & Kasneci, E. (2020). RemoteEye: An open-source high-speed remote eye tracker. *Behavior Research Methods, 52*, 1387–1401. doi:https://doi.org/10.3758/s13428-019-01305-2

[49] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., . . . Adam, H. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. Gjetur në arXiv:1704.04861

[50] Huang, Q., Veeraraghavan, A., & Sabharwal, A. (2015). TabletGaze: Unconstrained Appearance-based Gaze Estimation in Mobile Tablets. Gjetur në arXiv:1508.01244

[51] Hyllested, A., & Joseph, B. D. (2022). *The Indo-European Language Family, 13-Albanian.* (T. Olander, Re.) Cambridge University Press.

[52] Ivanchenko, D., Rifai, K., Hafed, Z. M., & Schaeffel, F. (2020). A low-cost, high-performance video-based binocular eye tracker for psychophysical research. *Journal of Eye Movement Research*. doi:10.16910/jemr.14.3.3

[53] Jakobi, D. N., Stegenwallner-Schütz, M., Hollenstein, N., Ding, C., Kaspere, R., Matic Skoric, A., . . . De Lopez, K. M. (2025). MultiplEYE: Creating a multilingual eye-tracking-while-reading corpus. *Proceedings of the 2025 Symposium on Eye Tracking Research and Applications (ETRA '25)* (fv. 1-11). New York, NY, USA: Association for Computing Machinery. doi:https://doi.org/10.1145/3715669.3726843

[54] Ji, X., Klehr, M., Ilieva, S., Rautenstrauch, J., & Franke, M. (2025). *Magpie Framework*. Gjetur në https://magpie-experiments.org/05_about/

[55] Kastrati, M., & Biba, M. (2022). Natural language processing for Albanian: a state-of-the-art survey. *International Journal of Electrical and Computer Engineering (IJECE), 12*(6), 6432-6439. doi:http://doi.org/10.11591/ijece.v12i6.pp6432-6439

[56] Ke, J., Jin, Q., You, S., Han, T., Feng, Y., Wang, X., . . . Ai, W. (2012). *The sqglobe corpus (a balanced corpus of 1m-word contemporary written albanian, lemmatised and pos-tagged).* Gjetur në http://114.251.154.212/cqp/sqglobe_v1/

[57] Kennedy, A., Hill, R., & Pynte, J. (2003). The Dundee Corpus. *Proc. of the 12th European Conference on Eye Movements.*

[58] Khatter, K., & Singh, S. (2022, 07). Natural language processing: state of the art, current trends and challenges. *Multimedia Tools and Applications, 82*, 3713-3744. doi:10.1007/s11042-022-13428-4

[59] Khonglah, J. R., & Khosla, A. (2015). A Low Cost Webcam Based Eye Tracker for communicating through the eyes of young children with ASD. *1st International*

Conference on Next Generation Computing Technologies (NGCT). Dehradun, India: IEEE. doi:10.1109/NGCT.2015.7375255

[60] Klerke, S., & Plank, B. (2019). At a Glance: The Impact of Gaze Aggregation Views on Syntactic Tagging. *Proceedings of the Beyond Vision and LANguage: inTEgrating Real-world kNowledge (LANTERN)* (fv. 51–61). Hong Kong, China: Association for Computational Linguistics. doi:10.18653/v1/D19-6408

[61] Kote, N., & Biba, M. (2022). Preprocessing tools for the Albanian language: A state of the art survey and an annotation schema proposal. *ALBANIAN JOURNAL OF NATURAL AND TECHNICAL SCIENCES (AJNTS), XXVII*(54). Gjetur në chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://akad.gov.al/wp-content/uploads/2024/02/01_Kote_Biba_AJNTS2022_1.pdf

[62] Kote, N., Biba, M., Kanerva, J., Rönnqvist, S., & Ginter, F. (2019). Morphological Tagging and Lemmatization of Albanian: A Manually Annotated Corpus and Neural Models. *arXiv*. Gjetur në https://arxiv.org/abs/1912.00991

[63] Kote, N., Rushiti, R., Çepani, A., Haveriku, A., Trandafili, E., Meçe, E. K., . . . Deda, A. (2024). Universal Dependencies Treebank for Standard Albanian: A New Approach. *Proceedings of the Sixth International Conference on Computational Linguistics in Bulgaria (CLIB 2024)* (fv. 80-89). Sofia, Bulgaria: Department of Computational Linguistics, Institute for Bulgarian Language, Bulgarian Academy of Sciences. Gjetur në https://aclanthology.org/2024.clib-1.7

[64] Krafka, K., Khosla, A., Kellnhofer, P., Kannan, H., Bhandarkar, S., Matusik, W., & Torralba, A. (2016). Eye Tracking for Everyone. *EEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016*. Gjetur në https://gazecapture.csail.mit.edu/

[65] Krakowczyk, D. G., Reich, D. R., Chwastek, J., Jakobi, D., Prasse, P., Suss, A., . . . Jager, L. A. (2023). pymovements: A Python Package for Eye Movement Data Processing. *ETRA '23: 2023 Symposium on Eye Tracking Research and Applications* (fv. 1 - 2). Tubingen, Germany: Association for Computing Machinery. doi:10.1145/3588015.3590134

[66] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM, 60*(6), 84-90. doi:https://doi.org/10.1145/3065386

[67] Laurinavichyute, A. K., Sekerina, I. A., Alexeeva, S., Bagfasaryan, K., & Kliegl, R. (2018). Russian Sentence Corpus: Benchmark measures of eye movements in reading in Russian. *Behavior Research Methods volume 51*, pages1161–1178. doi:https://doi.org/10.3758/s13428-018-1051-6

[68] Leal, S. E., Lukasova, K., Carthery-Goulart, M. T., & Aluisio, S. M. (2022). RastrOS Project: Natural Language Processing contributions to the development of an eye-tracking corpus with predictability norms for Brazilian Portuguese. *Language Resources and Evaluation*, 1333–1372. doi:https://doi.org/10.1007/s10579-022-09609-0

[69] Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE, 86*, 2278-2324. doi:10.1109/5.726791

[70] Lei, Y. (2021). Eye Tracking Calibration on Mobile Devices. *ETRA '21 Adjunct: ACM Symposium on Eye Tracking Research and Applications.* New York, NY, USA: Association for Computing Machinery. doi:https://doi.org/10.1145/3450341.3457989

[71] Lemhofer, K., & Broersma, M. (2011). Introducing LexTaLE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, 325-343. doi:10.3758/s13428-011-0146-0

[72] Lewandosky, S., Oberauer, K., Yang, L.-X., & K H Ecker, U. (2010). A working memory test battery for MATLAB. *Behav Res Methods*, 571-585. doi:10.3758/BRM.42.2.571

[73] Li, B., Snider, J. C., Wang, Q., Mehta, S., Foster, C., Barney, E., . . . Shic, F. (2022). Calibration Error Prediction: Ensuring High-Quality Mobile Eye-Tracking. *ETRA '22: 2022 Symposium on Eye Tracking Research and Applications* (fv. 1–7). New York, NY, USA: Association for Computing Machinery. doi:https://doi.org/10.1145/3517031.3529634

[74] Luke, S. G., & Christianson, K. (2017). The Provo Corpus: A large eye-tracking corpus with predictability norms. *Behavior Research Methods volume 50*. doi:https://doi.org/10.3758/s13428-017-0908-4

[75] Luo, X., Shen, J., Zeng, H., Song, A., Xu, B., Li, H., . . . Hu, C. (2019). Interested Object Detection based on Gaze using Low-cost Remote Eye Tracker. *9th International IEEE EMBS Conference on Neural Engineering* (fv. 1101-1104). San Francisco, CA, USA: IEEE. doi:10.1109/NER.2019.8716971

[76] McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics 17*, 578-586. doi:https://doi.org/10.3758/BF03203972

[77] Mishra, A., & Bhattacharyya, P. (2018). *Cognitively inspired Natural Language Processing: An Investigation based on Eye-tracking.* Springer Nature Singapore. doi:https://doi.org/10.1007/978-981-13-1516-9

[78] Mishra, A., Kanojia, D., Nagar, S., Dey, K., & Bhattacharyya, P. (2017). Leveraging Cognitive Features for Sentiment Analysi. *The SIGNLL Conference on Computational Natural Language Learning (CoNLL 2016).* Gjetur në https://arxiv.org/pdf/1701.05581.pdf

[79] Morozova, M., Rusakov, A., & Arkhangelskij, T. (2025, 01 12). *Korpusi Nacional i Gjuhës Shqipe*. Gjetur në https://albanian.web-corpora.net/index_sq.html

[80] Mounica, M. S., Manvita, M., Jyotsna, C., & J., A. (2019). Low Cost Eye Gaze Tracker Using Web Camera. *Proceedings of the Third International Conference on Computing Methodologies and Communication (ICCMC 2019)*, (fv. 79-85). doi:10.1109/ICCMC.2019.8819645

[81] Muller, D., & Mann, D. (2021). Algorithmic gaze classification for mobile eye-tracking. *ACM Symposium on Eye Tracking Research and Applications (ETRA '21 Adjunct).* New York, NY, USA: Association for Computing Machinery. doi:https://doi.org/10.1145/3450341.3458886

[82] MultiplEYE. (2024, Gusht 28). *Dominant Eye Test* . Gjetur në https://www.youtube.com/watch?v=dBCU6W2Pw6A&t=3s

[83] MultiplEYE-COST. (2025, Dhjetor 24). *WG1: Experiment Implementation*. Gjetur në Github: https://github.com/MultiplEYE-COST/wg1-experiment-implementation

[84] Nivre, J., de Marneffe, M., Ginter, F., Hajič, J., Manning, C., Pyysalo, S., . . . Zeman, D. (2020). Universal Dependencies v2: An Evergrowing Multilingual Treebank Collection. *Proceedings of the 12th Language Resources and Evaluation Conference* (fv. 4034–4043). Marseille, France: European Language Resources Association. Gjetur në https://aclanthology.org/2020.lrec-1.497/

[85] Papoutsaki, A., Laskey, J., & Huang, J. (2017). SearchGazer: Webcam Eye Tracking for Remote Studies of Web Search. *Proceedings of the ACM SIGIR Conference on Human Information Interaction \& Retrieval (CHIIR).* ACM. doi:https://doi.org/10.1145/3020165.3020170

[86] Papoutsaki, A., Sangkloy, P., Laskey, J., Daskalova, N., Huang, J., & Hays, J. (2016). WebGazer: ScalableWebcam Eye Tracking Using User Interactions. *IJCAI'16: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, 3839–3845.

[87] Park, J. H., & Park, J. B. (2016). A novel approach to the low cost real time eye mouse. *Computer Standards & Interfaces, 44*, 169-176. doi:https://doi.org/10.1016/j.csi.2015.04.005

[88] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., . . . Chintala, S. (2019). PyTorch: an imperative style, high-performance deep learning library.

Proceedings of the 33rd International Conference on Neural Information Processing Systems (fv. 8026 - 8037). Curran Associates Inc. doi:10.5555/3454287.3455008

[89] Pimsleur, P. (1966). Pimsleur Language Aptitude Battery (PLAB).

[90] Qi, P., Zhang, Y., Zhang, Y., Bolton, J., & Manning, C. D. (2020). Stanza: A Python Natural Language Processing Toolkit for Many Human Languages. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations.* Gjetur në https://nlp.stanford.edu/pubs/qi2020stanza.pdf

[91] Rakhmatulin, I., & Duchowski, A. T. (2020). Deep Neural Networks for Low-Cost Eye Tracking. *24th International Conference on Knowledge-Based and Intelligent Information & Engineering*. doi:https://doi.org/10.1016/j.procs.2020.09.041

[92] Raymond, O., Moldagali, Y., & Madi, N. A. (2023). A Dataset of Underrepresented Languages in Eye Tracking Research. *Proceedings of the 2023 Symposium on Eye Tracking Research and Applications (ETRA '23)* (fv. 1–2). Association for Computing Machinery. doi:https://doi.org/10.1145/3588015.3590128

[93] Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Q J Exp Psychol (Hove)*, 147-506. doi:10.1080/17470210902816461

[94] Rayner, K., Sereno, S. C., Morris, R. K., Scmauder, R., & Clifton, C. (2007). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes, 4*(3-4), S121-S149. doi:https://doi.org/10.1080/01690968908406362

[95] Ribeiro, T., Brandl, S., Søgaard, A., & Hollenstein, N. (2023). WebQAmGaze: A MultilingualWebcam Eye-Tracking-While-Reading Dataset. *Computation and Language (cs.CL)*. Gjetur në https://arxiv.org/pdf/2303.17876v2.pdf

[96] Rohanian, O., Taslimipoor, S., Yaneva, V., & Ha, L. A. (2017). Using Gaze Data to Predict Multiword Expressions. *Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP 2017* (fv. 601–609). Varna, Bulgaria: INCOMA Ltd. doi:10.26615/978-954-452-049-6_078

[97] Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. *ETRA '00*, 71–78. doi:10.1145/355017.355028

[98] Savitzky, A., & Golay, M. J. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry, 36*(8), 1627-1639. doi:10.1021/ac60214a047

[99] Saxena, S., Lange, E. B., & Fink, L. K. (2022). Towards efficient calibration for webcam eye-tracking in online experiments. *ETRA '22: 2022 Symposium on Eye Tracking Research and Applications* (fv. 1–7). Association for Computing Machinery. doi:https://doi.org/10.1145/3517031.3529645

[100] Scmidt, T., & Kai, W. (2014). *EXMARaLDA. Handbook of Corpus Phonology.* Oxford: Oxford University Press. Gjetur në www.exmaralda.org

[101] Shahid, A., Wilkinson, K., Marcu, S., & Shapiro, C. M. (a.d.). Karolinska Sleepiness Scale (KSS). Në *STOP, THAT and One Hundred Other Sleep Scales* (fv. 209-210). New York, NY.: Springer. doi:https://doi.org/10.1007/978-1-4419-9893-4_47

[102] Slegelman, N., Schroeder, S., Acaturk, C., Ahn, H.-D., Alexeeva, S., Amenta, S., . . . Kuperman, V. (2022). Expanding horizons of cross-linguistic research on reading: The Multilingual Eye-movement Corpus (MECO). *Behavior Research Methods*, 2843–2863. doi:https://doi.org/10.3758/s13428-021-01772-6

[103] SR Research Ltd. (2016-2025). *EyeLink Portable Duo User Manual.* Oakville, Ontario, Canada: SR Research. Gjetur në https://www.sr-research.com/support/attachment.php?aid=3078

[104] Strzyz, M., Vilares, D., & Gómez-Rodríguez, C. (2019). Towards Making a Dependency Parser See. *Computation and Language*. Gjetur në https://arxiv.org/pdf/1909.01053v1.pdf

[105] Taban, R. A., Croock, M. S., & Korial, A. E. (2018). Eye Tracking Based Mobile Application. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 7*(3).

[106] Talamo, L. (2025). Introducing STAF: The Saarbrücken Treebank of Albanian Fiction. *Journal of Open Humanities Data,, 11*(3), 1-6. doi:https://doi.org/10.5334/johd.285

[107] Tobii. (2025). Gjetur në Tobii: https://www.tobii.com/

[108] Toska, M., Nivre, J., & Zeman, D. (2020). Universal Dependencies for Albanian. *Proceedings of the Fourth Workshop on Universal Dependencies (UDW 2020)* (fv. 178–188). Barcelona, Spain (Online): Association for Computational Linguistics. Gjetur në https://aclanthology.org/2020.udw-1.20/

[109] Verkerk, A., & Talamo, L. (2024). mini-CIEP+ : A Shareable Parallel Corpus of Prose. *Proceedings of the 17th Workshop on Building and Using Comparable Corpora (BUCC) @ LREC-COLING 2024* (fv. 135–143). Torino, Italia: ELRA and ICCL. Gjetur në https://aclanthology.org/2024.bucc-1.15/

[110] Wang, Y., Zeng, H., & Liu, J. (2016). Low-cost eye-tracking glasses with real-time head. *10th International Conference on Sensing Technology (ICST)* (fv. 1-5). Nanjing, China: IEEE. doi:10.1109/ICSensT.2016.7796336

[111] Wilcox, E. G., Ding, C., Sachan, M., & Jäger, L. A. (2024). Mouse Tracking for Reading (MoTR): A new naturalistic incremental processing measurement tool.

Journal of Memory and Language, 138, 104534. doi:https://doi.org/10.1016/j.jml.2024.104534

[112] Wilcox, E. G., Ding, C., Sachan, M., & Jäger, L. A. (2024, October). Mouse Tracking for Reading (MoTR): A new naturalistic incremental processing measurement tool. *Journal of Memory and Language, 138*. doi:https://doi.org/10.1016/j.jml.2024.104534

[113] Wisiecka, K., Krejtz, K., Krejtz, I., Sromek, D., Cellary, A., Lewandowska, B., & Duchowski, A. T. (2022). Comparison ofWebcam and Remote Eye Tracking. *2022 Symposium on Eye Tracking Research and Applications*. doi:https://doi.org/10.1145/3517031.3529615

[114] Xu, P., Ehinger, K. A., Zhang, Y., Finkelstein, A., Kulkarni, S. R., & Xiao, J. (2015). TurkerGaze: Crowdsourcing Saliency with Webcam based Eye Tracking. Gjetur në arXiv:1504.06755

[115] Zhang, L., Wang, S., & Liu, B. (2018). ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices. *IEEE/ CVF Conference on computer Vision and Pattern Recognitin (CVPR)*, 6848-6856. doi:10.1109/CVPR.2018.00716

[116] Zhang, X., Park, S., Beeler, T., Bradley, D., Tang, S., & Hilliges, O. (2020). ETH-XGaze: A Large Scale Dataset for Gaze Estimation under Extreme Head Pose and Gaze Variation. *arXiv*. Gjetur në https://arxiv.org/pdf/2007.15837v1.pdf

[117] Zhang, X., Sugano, Y., Fritz, M., & Bulling, A. (2015). Appearance-Based Gaze Estimation in the Wild. *IEEE Conference on computer Vision and Pattern Recognition (CVPR)*, (fv. 4511-4520). Boston. doi:10.1109/CVPR.2015.7299081